CONVERGENCE OF REFLECTED DIFFUSIONS FROM INTERACTING PARTICLES

Brendan Brown

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Statistics and Operations Research.

Chapel Hill
2021

Approved by:

Sayan Banerjee

Shankar Bhamidi

Amarjit Budhiraja

Nicolas Fraiman

Vladas Pipiras

ABSTRACT

BRENDAN BROWN: Convergence questions for reflected diffusions arising from interacting particles
(Under the direction of Sayan Banerjee)

This work produces explicit convergence rates and properties of the stationary distributions for two different classes of reflected diffusions, each with an underlying representation as the interaction of stochastic particle systems. These are achieved by constructing explicit probabilistic couplings, as opposed to the implicit couplings of classical convergence techniques, and performing detailed analysis of their paths. We also take a step toward bringing this technique to questions of surrogate modeling and approximation of stochastic processes, which are typically seen in the context of engineering or applied math.

To the sperm whale who spied the ROV Hercules deep in the Gulf of Mexico, who taught me you can still make time for curiosity even when under great pressure.

`https://www.youtube.com/watch?v=SkBpummjR5I`

ACKNOWLEDGEMENTS

A warm thanks to those to whom I accumulate debt with joy:

CJ, SB and SB, AB, VP, NF, LJ, MB, KHB, JRB, LRSB, FJB, ORLS, CQ, MTN, ABJ, LRN.

TABLE OF CONTENTS

CHAPTER 1

Introduction

Stability and convergence questions for Markov processes range wide across probability and statistics, from shuffling cards Aldous and Diaconis (1986) to Bayesian statistics Hoff (2009, Ch. 10.4.2); from applied stochastic processes O'Connor et al. (2020); Biswas et al. (2019) to analysis Villani (2009). This work approaches the topic by studying certain classes of reflected diffusions in which relatively simple explicit probabilistic couplings, paired with a detailed analysis, bear explicit information about the stationary distribution or rates of convergences that classical techniques cannot. In that way, the work is best understood in the context of works such as Villani (2009); Eberle (2016), to give just a flavor, which have similar goals but for different objects treated with different techniques.

In this introduction, we give some context for the types of questions we will study in Chapters 2 and 3. Our focus is on the classical methods, which build intuition for how couplings play a role in convergence but prove to be insufficient, ultimately, for our goals here.

Chapter 2 looks at slightly weird class of reflected diffusions called inert drift systems. These are multidimensional reflected diffusions whose main feature is that the drift term of the first component is a stochastic process for which noise enters only through the 'local time' of the first component's interactions with the boundary. These in general provide examples of diffusions without regular transition densities, which is a non-started for many recipes proving recurrence and convergence.

Chapter 3 turns to the well-used and generically named Reflected Brownian motion, a Brownian motion with constant drift contrained to the positive orthant. This model arose from queuing theory Harrison and Reiman (1981) and remains, along with its generalizations, a topic of frequent study. Here we investigate whether subsystems of a high-dimensional system can converge more quickly than the full model, with a 'dimension-free' rate. We are motivated by some applications in stochastic finance where convergence of lower-dimensional systems is a relevant question, and more generally by 'dimension-free' convergence results in more computational work. Reflected Brownian Motions were recently shown to have badly dimension-dependent convergence rates, in typical cases.

Chapter 4 is different: It is a riff on the core ideas of the previous chapters, in which explicit couplings in turn give explicit information when used appropriately. In this chapter, however, we consider that meta-narrative for the purpose of studying approximations for stochastic processes, rather than their stability. We draw from the engineering and applied math literature on surrogate and multifidelity modeling. This is a preliminary look at how such tools might be of use in that context.

We prefer to define notation within the chapter where it is used, and otherwise to use standard notation throughout while giving clarification in special cases when needed. For example: $\mathcal{C}(\mathcal{S}_1, \mathcal{S}_2)$ is the space of continuous functions from one separable and complete (Polish) space $\mathcal{S}_1$ to another, $\mathcal{S}_2$, endowed with the standard metric. $\mathcal{C}^k(\mathbb{R}, \mathbb{R})$ is the space of continuous functions with kth order continuous derivatives, $\mathcal{C}_0^k(\mathbb{R}, \mathbb{R})$ the subset of those functions that vanish at infinity, etc.

1.1 Classical theory of stability for Markov processes

1.1.1 Doeblin's theorem and Markov chains

Stability of Markov processes usually starts with the following fundamental convergence theorem, often called Doeblin's theorem, to which most of the ideas in the field can in some way be traced. We state Theorem 4.9 from Levin, D. and Peres, Y. and Wilmer, E. (2017) as Theorem 1.1.2 below and refer the reader to Chapters 4, 5 therein for a clean and illustrative treatment of the related theory.

Recall first the definition of total variation distance between two probability measures μ, ν on a countable space Ω as $\|\mu - \nu\|_{TV} = \sup_{A \subset \Omega} |\mu(A) - \nu(A)| = \frac{1}{2} \sum_{i \in \Omega} |\mu(i) - \nu(i)|$. An alternative characterization of the total variation distance that will help in illuminating Theorem 1.1.2 is,

$$\|\mu - \nu\|_{TV} = \inf \left\{ \sum_{i,j \in \Omega} \gamma(i,j) \, \mathbb{1} \{i \neq j\} \quad : \gamma \text{ is a coupling of } \mu, \nu \right\}, \tag{1.1.1}$$

and we give here the definition of probabilistic couplings in a more general setting, for future reference:

Definition 1.1.1 (Coupling). *A coupling γ of two probability measures μ, ν on an arbitrary measurable spaces $(\Omega_i, \mathcal{F}_i), i = 1, 2$ is a probability measure on the product space $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ such that*

$$\int_{A_1 \times \Omega_2} d\gamma = \mu(A_1), \qquad \int_{\Omega_1 \times A_2} d\gamma = \nu(A_2)$$

for all $A_i \in \mathcal{F}_i$.

Theorem 1.1.2. *Suppose X is a Markov chain with finite state space $\{1 \ldots N\}$ and transition matrix P, which is irreducible and aperiodic. Then, there exists a unique stationary distribution π such that for any initial condition i the transition laws $P^n(i, \cdot)$ converge to π exponentially fast in total variation.*

Specifically, there exists constants $\lambda \in (0, 1), K > 0$ such that

$$\sup_{1 \leq i \leq N} \|P^n(i, \cdot) - \pi\|_{TV} \leq K\lambda^n, \quad n \geq 1. \tag{1.1.2}$$

We quickly recall a few definitions used in the theorem. Suppose P is the transition matrix of a Markov chain $X = \{X_n\}_{n \geq 1}$ with state space $\{1 \ldots N\}$. X is called irreducible if the matrix P is irreducible, that is if for each i, j indices there exists an $m(i, j)$ such that $P^{m(i,j)}(i, j) > 0$. It is aperiodic if for all states i the greatest common divisor of the set $\{m \geq 1 : P^m(i, i) > 0\}$ is 1.

Theorem 1.1.2 is pleasing as can be: The key fact is that the assumptions provide a constant $\alpha > 0$ and step size m such that the transition laws P^m satisfy the *minorization condition*

$$P^m(i, j) \geq \alpha\pi(j), \qquad \text{for all } i, j \in \{1 \ldots N\}. \tag{1.1.3}$$

Conceptually, the theorem's proof proceeds as follows: For simplicity suppose $m = 1$. Construct a coupling of two versions of X. The first, X^π has initial law π. For the second, at each time step with probability α (independently of both processes) set it equal X^π and run the two processes together from then on. Otherwise, sample its transition according to

$(P-\alpha\pi)/(1-\alpha)$. This forms a coupling such that, independently for each step, the probability the two chains are not equal is α. The theorem now follows with $\lambda = 1-\alpha$ by (1.1.1) and the fact that $n \mapsto \|P^n(i,\cdot)-\pi\|_{TV}$ is contractive.

Finite-state Markov chains are among the special cases of Markov processes in which there is some hope for the rate of convergence to stationarity to be explicit, for example when the matrix P is lower bounded.

For countable-state Markov chains, which again are irreducible and aperiodic, many of the same ideas can be applied to achieve similar results if the chain is *positive recurrent*, meaning the expected return time to any state from which it started is finite. In this case, it is the moments of the return time that determine rate of convergence on the right-hand side of (1.1.2). For example, Pitman (1974) shows finite polynomial moments of return times produce polynomial rates of convergence. However, instead of reviewing such results in any detail, we skip instead to systems more similar to those studied here.

1.1.2 Probabilistic convergence methods for general Markov processes

Here we consider time-homogeneous Markov processes indexed by $[0,\infty)$ whose state space is a complete, separable metric (Polish) space $\mathcal{S}$, and to avoid needless complexity we take $\mathcal{S} \subset \mathbb{R}^d$ throughout this section. The main ingredients leading to Theorem 1.1.2 are present here, roughly speaking, but take on a greater technical burden. We review conditions under which we can recover exponentially fast convergence using classical generator methods.

A time-homogeneous Markov process X is defined by a family of transition probability kernels $\{P_t\}$ whose key feature is the Markov property,

$$\mathbb{E}\left(f(X_{t+s}) \,|\, \mathcal{F}_s\right) = P_t f(X_s), \qquad f \in \mathcal{B}(\mathcal{S}), \tag{1.1.4}$$

where $\mathcal{B}(\mathcal{S})$ denotes the bounded Borel-measurable functions on $\mathcal{S}$, $\{\mathcal{F}_s\}_{s\geq 0}$ the filtration to which X is adapted and we use the notation

$$P_t f(x) = \int_{\mathcal{S}} f(y) P_t(x, dy). \tag{1.1.5}$$

By definition, $x \mapsto P_t f(x)$ is measurable for each $f \in \mathcal{B}(\mathcal{S})$ and for each x the map $A \mapsto P_t \mathbb{1}_A(x)$ is a probability measure. For formal, complete definitions of transition probability kernels we refer the reader to Le Gall (2013); Revuz and Yor (1999) or for an analytic treatment via the Hille-Yoside theorem Evans (2010, Ch. 7.4).

Notions of irreducibility and recurrence now require some technical modifications to the concepts of Section 1.1 but in broad strokes remain the same Meyn and Tweedie (1993a): Irreducibility holds when the process must spend some interval of time, with positive probability, in any set with positive mass with respect to a reference measure, usually Lebesgue measure. Here (Harris) recurrence holds if the hitting time of the process is a.s. finite for sets of positive mass with respect to the reference, for every initial condition of the process. This is enough to guarantee a measure invariant under $\{P_t\}_{t\geq 0}$, and if this measure is finite the process is said to be positive (Harris) recurrent. Positive recurrence, along with a slightly more technical irreducibility condition, are enough to show total variation convergence Meyn and Tweedie (1993a, Thm. 6.1), but without a rate as in Theorem 1.1.2.

Two ingredients are needed to achieve exponential rates of convergence. First is a condition like (1.1.3) that ensures there is a reachable subset of the state-space from the marginals in a coupling of the process can meet with positive probability. This states that there exists a $t_0 > 0$, a probability measure μ and $\alpha \in (0,1)$ such that

$$P_{t_0}(x, \,\cdot\,) \geq \alpha\mu(\cdot) \quad x \in \Lambda, \qquad \textit{(Minorization)} \tag{1.1.6}$$

where Λ is an appropriate compact set on which there are some technical conditions too lengthy to state here. See the discussion surrounding (2.6.23) below.

Unlike in the finite-state case, however, we must ensure here that the process returns to the set Λ where coupling can occur sufficiently quickly. This requirement is codified in the following 'drift condition:' There exists a function $V \geq 0$ and constants $t_1, K > 0$, $\lambda \in (0,1)$ such that

$$P_{t_1}V(x) \leq \lambda V(x) + K \quad x \in \mathcal{S}, \qquad \textit{(Drift)} \tag{1.1.7}$$

where V is typically called the Lyapunov function. The constant λ determines the rate of convergence as in Theorem 1.1.2, and (1.1.7) can be relaxed to hold on a set Λ like that in 1.1.6. A fact helping to demystify (1.1.7) is that such V can be constructed via exponential moments of the return times to sets meeting the requirements for Λ in (1.1.6). See Down et al. (1995); Meyn and Tweedie (2009); Hairer and Mattingly (2011) and Lemma 2.6.4 here.

The minorization and drift conditions are conceptually simple but can be technically difficult to verify for general Markov processes. The conditions have bubbled up from the discrete-time, discrete-space Markov chain theory and thus are easiest to apply when $(t,x) \mapsto P_t f(x)$ is sufficiently regular for bounded measurable f. We turn to this nice scenario now, before discussing the more problematic processes studied in Chapters 2 and 3.

We make the final comment that the recipe encoded in (1.1.6), (1.1.7) is very robust despite the technical challenges it poses in general scenarios. For a small sampling of the breadth of this approach refer to Cloez and Hairer (2015); Butkovsky et al. (2020); Budhiraja and Lee (2007); Bertoin (2019), Thorisson (2000, Ch. 3.6) in the context of 'spread-out' random walks, and Chapter 2.

1.1.3 Diffusions with smooth transition densities

In this section we consider Markov processes in which the transition laws $\{P_t\}_{t>0}$ admit sufficiently regular densities. We review the conditions under which this holds in the context of diffusions, the study of which forms the core of this work, and briefly describe how regularity of densities relieves some of the abstraction and technical burden of (1.1.6) when proving exponential convergence to stationarity.

Here we view $\{P_t\}_{t\geq 0}$ as a family of operators on $\mathcal{C}_0(\mathcal{S})$, the real-valued functions vanishing at infinity. $\{P_t\}_{t\geq 0}$ is said to be (weak) Feller (or equivalently X is Feller) if $x \mapsto P_t f(x)$ is in $\mathcal{C}_0(\mathcal{S})$ for each $f \in \mathcal{C}_0(\mathcal{S})$ and for each f, x it holds that $\lim_{t\downarrow 0} P_t f(x) = f(x)$ Revuz and Yor (1999, Prop. 2.4). In this case, we define the infinitesimal generator as

$$\mathcal{L}f(x) = \lim_{t\downarrow 0} \frac{P_t f(x) - f(x)}{t}, \qquad f \in \mathcal{D}(\mathcal{L}), \tag{1.1.8}$$

$\mathcal{D}(\mathcal{L})$ is the domain in $\mathcal{C}_0(\mathcal{S})$ where the limit exists for each x. $\mathcal{L}$ and its domain determine $\{P_t\}_{t\geq 0}$ Le Gall (2013, Cor. 6.13).

(1.1.7) can be recast in terms of $\mathcal{L}$ in a straightforward way. This section, though, will focus on the classic method by which study of $\mathcal{L}$ can overcome the technical difficulties imposed by the minorization condition (1.1.6). Of particular relevance are the generators given as second-order differential operators, which correspond to diffusions.

Definition 1.1.3 (Diffusion). *A diffusion is a time-homogeneous Markov process $X \in \mathcal{C}\left([0,\infty),\mathbb{R}^d\right)$ that is the solution to the Itô stochastic differential equation*

$$dX_t^x = b(X_t^x)dt + \sigma(X_t^x)dW_t, \tag{1.1.9}$$

with $X_0 = x$, where W is a d-dimensional Brownian motion, and $\sigma : \mathbb{R}^d \mapsto Hom(\mathbb{R}^d, \mathbb{R}^d)$ the $d \times d$ matrices, $b : \mathbb{R}^d \mapsto \mathbb{R}^d$ are assumed to be Lipschitz continuous in the respective standard norms. The operator $P_t f(x) := \mathbb{E} f(X_t^x)$ has generator,

$$\mathcal{L}f(x) = \sum_{1}^{d} b_i(x)\partial_{x_i} f(x) + \frac{1}{2}\sum_{i,j=1}^{d} \left(\sigma\sigma^T\right)_{ij}(x)\partial_{x_i x_j} f(x), \tag{1.1.10}$$

for $f \in \mathcal{C}_c^2(\mathbb{R}^d) \subset \mathcal{D}(\mathcal{L})$ the twice continuously differentiable functions of compact support.

In a standard application we would also assume the following non-degeneracy condition on the covariance matrix $\sigma\sigma^T$.

Definition 1.1.4 (Ellipticity). *A matrix-valued function $a : \mathbb{R}^d \mapsto Hom(\mathbb{R}^d, \mathbb{R}^d)$ is called elliptic if there exists a constant $c > 0$ such that*

$$\sum_{i,j=1}^{d} a_{ij}(x)x_i x_j \geq c\|x\|^2, \quad x \in \mathbb{R}^d \tag{1.1.11}$$

where $\|\cdot\|$ is the standard Euclidean norm. We call a generator $\mathcal{L}$ of the form (1.1.10) *elliptic if the covariance matrix $\sigma\sigma^T$ satisfies* (1.1.11).

Return now to the key minorization condition for convergence given in (1.1.6). If $\mathcal{L}$ is elliptic and $\sigma\sigma^T, b$ are smooth, meaning with continuous derivatives of all orders, and if these derivatives are bounded, then by classical theory of elliptic second-order differential operators the transition laws $\{P_t\}$ have densities $p_t(x, y)$ that are smooth in (t, x, y) Stroock (2008, Thm. 3.4.1). See also Evans (2010, Ch. 6.3) for results on weak solutions to elliptic PDEs in cases where $\sigma\sigma^T, b$ have less regularity. In fact, so long as (1.1.11) holds, if $\sigma\sigma^T, b$ are twice continuously differentiable with bounded derivatives then there exists densities $p_t(x, y)$ that are continuously differentiable in t and twice continuously differentiable in y. See Friedman (1975, Thm. 4.7 Ch. 6, and discussion following Thm. 5.4) and Nualart (2006, Thm. 2.3.1) for existence of densities under weaker conditions, with no regularity guarantees.

In cases where (1.1.11) fails we can nonetheless recover smooth transition densities if we can show that the operator $\mathcal{L}^* - \partial_t$ is *hypoelliptic*, where $\mathcal{L}^*$ is the adjoint of $\mathcal{L}$. Hypoellipticity means that if u is a distribution on the smooth functions of compact support $\mathcal{C}_c^\infty$, then if $(\mathcal{L}^* - \partial_t)\, u$ is smooth u itself must be smooth (where derivatives of the distribution are defined in a weak sense). In that case, we use the fact $(\mathcal{L}^* - \partial_t)\, P_t = 0$ (in a weak sense) to determine P_t has smooth transition densities. Since such methods are out of scope for this work, we defer to Stroock (2008, Thm. 7.4.20) for applications to diffusions such as (1.1.9) and to Simon (2015, Ch. 6) for basics on distribution theory.

When sufficiently regular transition densities exist, proving (1.1.6) reduces to showing that there exists a compact set Λ, a point $x_0 \in \Lambda$ and a time $t_0 > 0$ such that the transition density $p_{t_0}(x_0, \cdot)$ is strictly positive in some neighborhood contained in Λ, and such that any point in Λ can reach a neighborhood of x_0 with strictly positive chance.

This is a substantial reduction in the technicality and abstraction inherent in (1.1.6), which is why transition densities play such an important role in the classical theory of exponentially fast convergence results akin to Theorem 1.1.2.

For a succinct and concrete application of this procedure with a minimum of regularity conditions imposed on the transition density, see Mattingly et al. (2012, Prop. 2.4). For a sampling of other work where this method is exploited, see Mattingly and Stuart (2002); Cooke et al. (2017); Athreya et al. (2012); Herzog and Mattingly (2015). For a clean exposition of the connections between exponential moments of return times, Lyapunov conditions such as (1.1.7) and other stability methods such as the Poincaré inequality, see Cattiaux and Guillin (2017).

1.2 Diffusions with reflection arising from interacting particle systems

In both Chapters 2 and 3 we study diffusions with reflecting boundary conditions, in which the reflections represent interactions between particles in some underlying process. Since we will describe the theory of these processes in some detail in subsequent chapters, in this introduction we focus on just two messages:

First, reflected diffusions disrupt the main tools needed to prove exponential convergence in Section 1.1. Second, resolving these difficulties pushes us toward techniques that shine where classical methods typically fail, using constructive couplings that produce more explicit rates of convergence. This places our work in a broader stream of research to derive explicit convergence resultsfor both theoretical and statistical purposes, for example as applied to Monte Carlo algorithms.

A generic reflected diffusion can be written as a solution to a constrained version of (1.1.9),

$$dX_t^x = b(X_t^x)dt + \sigma(X_t^x)dB_t + R(X_t^x)dL_t^x. \tag{1.2.1}$$

Here, L_t^x is a constraining process guaranteeing X^x remains in some domain with boundary $\mathcal{S}_0 \subset \mathbb{R}^d$ for all time $x \in \mathcal{S}_0$, and R is a matrix determining the direction of reflection when X^x makes contact with the boundary. L is in typical cases non-decreasing, right-continuous and constant for t such that X_t^x is in the interior of the domain. Unlike for (1.1.9), basic existence and uniqueness questions for these diffusions in generality are not straightforward and instead require methods sensitive to the domain $\mathcal{S}_0$ and the reflections $R(X_t^x)$ Lions and Sznitman (1984); Knight (2001, e.g.).

When it comes to convergence and stability, we are concerned with two types of problems: First are cases when the diffusion matrix $\sigma\sigma^T$ is not elliptic in the sense of Definition 1.1.4. Secondly, even under conditions where classical convergence methods can be used, the lack of explicit rates of convergence they provide are particularly concerning for reflected systems, where non-smooth boundary interactions can slow convergence considerably. We discuss these problems in turn.

1.2.1 Lack of transition densities

As discussed in the previous section, degeneracy robs us of a hugely useful tool in proving recurrence and convergence. But in the unreflected case there is more hope that, if not elliptic, the process nonetheless is hypoelliptic and hence accrues the benefits of smooth transition densities.

A simple example of a non-reflected system where this occurs is the diffusion with generator $\mathcal{L}f(x) = \frac{1}{2}\partial_{x_1x_1}f(x) + x_1\partial_{x_2}f(x)$ Stroock (2008, Ch. 7.4, who attributes the example to Kolmogorov). The diffusion matrix $\sigma\sigma^T$ has 1 in the top-left entry and zeros otherwise. Its associated Markov process has as a first coordinate $X_t^x = x_1 + B_t$ and as the second coordinate $X_t^x = x_2 + tx_1 + \int_0^t B_s ds$. This is a bivariate, non-degenerate Gaussian process which therefore has a smooth transition density for $t > 0$, and in fact the generator is hypoelliptic.

In a way this example is a dividing line between the previous sections and the next. The generator of the process considered in Chapter 2 is similarly degenerate, but in the latter case the noise enters the second coordinate only via boundary interactions, which occur on a set of times with zero Lebesgue measure. Away from the boundary, the process deterministic movement in one coordinate and simple examples show a transition density with respect to Lebesgue measure cannot exist.

Degeneracy imposes greater technical challenges to existence and uniqueness of stationary distributions, and to proving exponentially fast convergence. But it lends itself nicely to a more detailed study of the paths. Deterministic motions in sections of the state space allow for greater control. Bass et al. (2010) gives a convenient example for this chapter, in Section 6 to prove irreducibility and in Proposition 4.8: The authors study diffusion constrained to a bounded domain with smooth boundary, with a drift term whose velocity is deterministic but for the influence of the 'local time' interactions of the process with the boundary. This is an 'inert drift' system, though in a somewhat different formulation that the one studied in Chapter 2. In Proposition 4.8, they use a pathwise analysis to show that, if one assumes a 'local' hypoellipticity at a special point, then that point must be in the support of any stationary distributionr, which in turn proves uniqueness. Similarly, in Section 6 they use a excursion-based analysis to show fundamental irreducibility properties for the system.

With methods that are similar at least in spirit, we establish in Chapter 2 some upper and lower tail bounds on the stationary distribution, along with long-time scaling limits, that are explicit in the system parameters. This is possible because of, not in spite of, the fact that noise travels in only one of two directions. The cited chapter contains multiple additional examples of work following this general line of attack.

1.2.2 Inexplicit rates from classical methods

Many reflected systems, however, do have nice regularity properties like those of Section 1.1.3. We can define a version of the generator in (1.1.10) for the reflected system (1.2.1), which takes the form (1.1.10) for x in the interior of the domain and additional boundary constraints. Studying these generators in cases where $\mathcal{S}_0$ is a half-space or is has sufficiently regular boundary, Ramasubramanian (1996) shows there exist positive transition densities for (1.2.1) under ellipticity, regularity conditions in σ, b, R, and a sufficiently inward-pointing angle of reflection in R. See also Harrison and Williams (1987b); Atar et al. (2001, e.g.) for other cases of analysis for reflected systems with good regularity in the transitions.

More relevant to the work of Chapter 3 are studies of the so-called Reflected Brownian Motion in the positive orthant, where the drift, diffusion and reflection coefficients in (1.2.1) are all constant. These systems are much-studied in the operations research literature in particular as they arise as limits of certain queueing systems Harrison and Reiman (1981, for the initial description in this context).

Using the regularity of reflected Brownian motion's transition laws, Budhiraja and Lee (2007) establish its exponentially fast convergence using the formula (1.1.6), (1.1.7). These rates are not explicit, however, and that is typical for the method. See also Theorem 2.1.6 and Tang (2019).

Essentially the problem is that (1.1.6), (1.1.7) contain too many existence statements best proved by non-constructive means. The underlying coupling procedure (discussed in that section) is quite implicit. Towers of abstraction, in fact, are baked into the proofs of the major works in this discipline Meyn and Tweedie (2009, 1993b); Down et al. (1995) including one of the only uses of Egorov's theorem in probability outside of a textbook this author has seen Meyn and Tweedie (1993a, Prop. 6.1).

This is vexing in cases where one wishes to study high-demensional systems, in particular. For example when dimension-explicit rates are computed for the reflected Brownian motion, as in Banerjee and Budhiraja (2020), we see that under strong assumptions the rate parameter of exponential convergence (meaning the r in $\exp(-rt)$) decreases like $1/\log d$ but under weaker assumptions decays like $1/d^4 \log d$ — which the abstract results of the previous section give no information about.

Hairer and Mattingly (2011) do make substantial improvements in this regard to the theory from Section 1.1.2, by tying the set Λ in the minorization condition to the Lyapunov function and the convergence rate explicitly to λ in the drift condition. However, the coefficient λ in the drift condition itself is usually difficult to make explicit.

We defer discussion of recent developments deriving explicit rates of convergence for diffusions to Chapter 3, where it is most relevant. The story in broad strokes is that by constructing couplings tied to the dynamics of the diffusion we may apply Lyapunov- and super-martingale type methods to achieve convergence rates explicit in system dimension or dynamics. The cost of doing so is to change the distance in which convergence is measured, from total variation distance (see the definition prior to Theorem 1.1.2) to various Wasserstein distances.

CHAPTER 2
Inert drift in a viscous fluid

2.1 Introduction and summary of results

We study an inert drift system which models the joint motion of a massive (inert) particle and a Brownian particle in a viscous fluid in the presence of a gravitational field. The inert particle is impinged from below by the Brownian particle which transfers momentum to it proportional to the 'local time' of collisions. This acceleration is countered by the viscosity of the fluid and the gravitational force acting on the inert particle.

The interaction is non-standard because transfer of momentum is 'non-Newtonian:' It can be thought of as an 'infinite number of collisions' between the particles in any finite time interval, such that each collision results in an 'infinitesimal momentum transfer.' This inert drift system is a simplified mathematical model for the motion of a semi-permeable membrane in a viscous fluid in the presence of gravity Einstein (1956); Knight (2001). The membrane plays the role of the inert particle, which is permeable to the microscopic fluid molecules but not the macroscopic Brownian particle. The following system of stochastic differential equations characterizes the joint motion of particles in the present model:

$$\begin{cases} dX_t = dB_t - dL_t \\ dV_t = -(\gamma V_t + g)dt + dL_t \\ dS_t = V_t dt \end{cases}$$

starting from $(X_0, V_0, S_0) \in \mathbb{R}^3$ with $S_0 \geq X_0$ and $V_0 \in (-g/\gamma, \infty)$. See (2.1.3) for justification of the restricted range for V_0.

Here $\gamma, g > 0$ respectively denote the viscosity coefficient and the acceleration due to gravity, B_t is a one-dimensional standard Brownian motion, S_t is the trajectory of the inert particle, V_t

is the velocity of the inert particle and L_t is the local time of collisions, which is defined as the unique continuous, non-negative, non-decreasing process such that $\int_0^t \mathbb{1}_{S_u - X_u = 0} dL_u = L_t$ and $S_t - X_t \geq 0$ for all t. We will assume $S_0 = 0$ unless otherwise stated.

Knight (2001) initiated the study of inert drift systems by studying the case where $g = \gamma = 0$. In this case, the system becomes transient, meaning the inert particle escapes. Knight (2001) determined the laws of the inverse velocity process and the 'escape velocity.' Since then, numerous inert drift systems have been studied Bass et al. (2010); White (2007); Burdzy and White (2008). Barnes (2018) studied hydrodynamic limits for inert drift type particle systems. Moreover, stochastic differential equations somewhat similar in flavor to inert drift systems have recently appeared as diffusion limits of queuing networks like the join-the-shortest-queue discipline Eschenfeldt and Gamarnik (2018); Banerjee and Mukherjee (2019a,b). Banerjee et al. (2019) studied the inert drift model with $g > 0, \gamma = 0$. Unlike Knight (2001) the model is recurrent, in other words the inert particle does not escape. The paper showed that the process hits a specified point almost surely and that the hitting times at this point has finite expectation. By decomposing the path into excursions between successive hitting times of such a point, the paper showed the process $(S_t - X_t, V_t)$ has a renewal structure and converges in total variation distance to a unique stationary distribution. Moreover, it was shown by solving an associated partial differential equation that the process $(S_t - X_t, V_t)$ has an explicit product form stationary distribution which is Exponential in the first coordinate and Gaussian in the second coordinate. Using this explicit form along with the renewal structure, the paper obtained sharp fluctuation estimates for the velocity V_t and the gap between the particles $S_t - X_t$.

In this chapter, we analyze the full model, in which $g > 0, \gamma > 0$. In contrast with Banerjee et al. (2019), there is no explicit closed form for the joint stationary distribution of the velocity and the gap. Nevertheless, we establish a renewal structure analogous to Banerjee et al. (2019) and thus obtain a tractable renewal theoretic representation of the stationary distribution (Theorem 2.1.2). By analyzing the excursions between successive renewal times, we obtain precise upper and lower bounds on the tails of the stationary distribution (Theorem 2.1.3). Besides furnishing Exponential tails for the gap and Gaussian tails for the velocity, these bounds display the explicit dependence of the stationary distribution tails on the model parameters g, γ. We exploit renewal structure further to obtain parameter-dependent fluctuation estimates for the gap $S_t - X_t$ and the velocity V_t (Theorem 2.1.4), which also imply law of large numbers results for S_t and X_t (Theorem 2.1.5).

One surprising aspect of the model arises from Theorem 2.1.4 which shows that $(S_t - X_t)/\log t$ is $O(\gamma/g)$ for large t as compared to $O(1/g)$ for the model without viscosity studied in Banerjee et al. (2019). Thus, the fluctuation results and tail estimates for the case $\gamma = 0$ cannot be recovered by taking $\gamma \to 0$ in our results. This shows that the *qualitative behavior of the steady state changes on introducing viscosity*, and that the rare events contributing to tail estimates of the stationary measure arise in very different ways between the $\gamma = 0$ and $\gamma > 0$ cases (see Remark 2.1.1 after Theorem 2.1.4).

In addition, we show that convergence to stationarity is exponentially fast, by appealing to Harris' Theorem (see Hairer and Mattingly (2011); Meyn and Tweedie (2009) for versions of the theorem). The standard approach to Harris' Theorem relies on the existence of continuous densities for transition laws with respect to Lebesgue measure (see Budhiraja and Lee (2007); Cooke et al. (2017); Mattingly et al. (2012); Cattiaux and Guillin (2017)) which, in turn, is shown by establishing that the generator of the process is hypoelliptic. However, the generator of our process $(S_t - X_t, V_t)$ is not hypoelliptic in the interior of the domain (when $S_t - X_t > 0$) and the transition laws of $(S_t - X_t, V_t)$ do not have densities with respect to Lebesgue measure. This fact is in essence a consequence of the velocity's deterministic evolution when $S_t - X_t > 0$. Therefore, possible ergodicity arises from non-trivial interactions between the drift, the Brownian motion and the boundary reflections.

To circumvent these technical challenges, we once again use the renewal structure of our process to show a minorization condition (2.6.23) and to obtain exponential moment estimates for hitting times to certain sets. Exponential moments provide a suitable Lyapunov function and thereby a drift condition (2.6.24) used to obtain exponential ergodicity via the techniques of Meyn and Tweedie (2009); Thorisson (2000); Down et al. (1995).

Now we give an outline of the organization of the article. In Subsections 2.1.3 and 2.1.4, we describe the renewal structure of the system and state the main results of the article. In Section 2.2 we prove the existence and uniqueness of the process and, in particular, prove its strong Markov property and a Skorohod representation for the local time. In Section 2.3, we obtain tail estimates on the distribution of the renewal time which, in particular, imply its integrability and the existence and uniqueness of the stationary measure. It also gives a tractable renewal theoretic representation of the stationary measure. In Section 2.4, fluctuation bounds for the velocity and gap process between two successive renewal times are obtained. These bounds imply tail estimates on the stationary measure and path fluctuation results, which are proved in Section 2.5. In Section 2.6, we prove that the process converges to its stationary measure exponentially fast in total variation distance. Finally, Appendix 4.4 at the end of this document is devoted to some technical estimates for hitting times which are used throughout the article.

2.1.1 Chapter notation

We set the following notation:

$$
\begin{aligned}
H_t &= S_t - X_t \\
\sigma(t) &= \inf\{s \geq t \mid H_s = 0\} \\
\tau_b^V &= \inf\{s \geq 0 \mid V_s = b\} \\
\tau_B^V &= \inf\{s \geq 0 \mid V_s \in B\} \\
\tau_x^H &= \inf\{s \geq 0 \mid H_s = x\} \\
\tau_A &= \inf\{s \geq 0 \mid (H_s, V_s) \in A\} \\
P^t\left((h,\nu), \cdot\,\right) &= \mathbb{P}_{(h,\nu)}\left((H_t, V_t) \in \cdot\,\right) = \mathbb{P}\left((H_t, V_t) \in \cdot \mid (H_0, V_0) = (h,\nu)\right) \\
P^t\left((h,\nu),\, f\right) &= \mathbb{E}_{(h,\nu)} f\left(H_t, V_t\right)
\end{aligned}
$$

where f is a measurable function such that $\mathbb{E}_{(h,\nu)} f\left(H_t, V_t\right)$ is defined, $b \in \mathbb{R}$, $x \geq 0$ and $B \subset \mathbb{R}, A \subset \mathbb{R}^2$ are Borel-measurable sets. Unless otherwise stated, $S_0 = X_0 = B_0 = 0$. The notation $|\cdot|$ will be used for the Euclidean norm on $\mathbb{R}$ or $\mathbb{R}^2$, the space being clear from context. We will work with system equations reformulated as

$$
\begin{cases}
dH_t = V_t\, dt - dB_t + dL_t \\
dV_t = -(\gamma V_t + g)\, dt + dL_t
\end{cases}
\tag{2.1.1}
$$

for initial conditions $(H_0, V_0) = (h, \nu) \in \mathbb{R}_+ \times \left(-\frac{g}{\gamma}, \infty\right)$ and local time L defined as the unique continuous, non-negative, non-decreasing process such that $\int_0^t \mathbb{1}_{H_t=0} dL_u = L_t$ and $H_t \geq 0$ for all t. Sometimes, we will write $L_t^{(h,\nu)}$ in place of L_t to elucidate dependence on the initial conditions (h, ν). We will show in Theorem 2.2.1 that the Skorohod representation for the local time is valid, namely

$$L_t = \sup_{u \leq t} \left(-h + B_u - \int_0^u V_w dw\right) \vee 0. \tag{2.1.2}$$

We assume $C' > 0$ are fixed throughout. We will write the state-space of the solution to (2.1.1) as $S = \mathbb{R}_+ \times \left(-\frac{g}{\gamma}, \infty\right)$, but also use S for the state-space of the more general diffusion (2.2.1) when there is no danger of confusion.

K, K' etc. will always denote positive constants, not depending on γ, g. c, c', C, C' etc. will denote positive constants dependent on γ, g. Values of constants might change from equation to equation without mention. Throughout, $\gamma, g > 0$ are fixed.

2.1.2 System properties

We adapt general techniques from stochastic differential equations to our context to show that a strong solution to (2.1.1) exists, is pathwise unique and has the strong Markov property, proven in Theorem 2.2.1.

We state a few fundamental properties of the system's motion that are integral to all results in this paper. For initial conditions $(H_0, V_0) = (h, \nu) \in S$, since L_t is non-negative for each $t \geq 0$, (2.1.1) shows

$$V_t \geq (v + g/\gamma)\, e^{-\gamma t} - g/\gamma > -g/\gamma \quad \forall\, t \geq 0, \tag{2.1.3}$$

which justifies the state-space S rather than $\mathbb{R}_+ \times \mathbb{R}$. Moreover, for $t < \sigma(0)$,

$$\begin{aligned}
V_t &= \nu - \int_0^t (\gamma V_s + g)\, ds = \nu - \gamma S_t - gt = (\nu + g/\gamma)\, e^{-\gamma t} - g/\gamma, \\
H_t &= h + S_t - B_t = h + \nu/\gamma - V_t/\gamma - B_t - tg/\gamma \\
&\leq h + \nu/\gamma + g/\gamma^2 - B_t - tg/\gamma
\end{aligned} \tag{2.1.4}$$

where the last inequality follows from $V_t \geq -g/\gamma$ for all $t \geq 0$. Thus H is dominated by Brownian motion with drift $-g/\gamma$ in S°. The last equality in the velocity equation in (2.1.4) comes from solving the ODE for V obtained from (2.1.1) without L_t, which is zero for $t \in [0, \sigma(0))$. (2.1.4) shows the velocity increases on ∂S and only there, i.e. at t when $H_t = 0$. Otherwise stated, if $(H_0, V_0) \in S^\circ$,

$$V_t \text{ is decreasing on } \quad t \in [0, \sigma(0)) \quad \text{and} \quad H_{\tau_b^V} = 0 \quad \text{for } b > V_0. \tag{2.1.5}$$

Properties (2.1.4) and (2.1.5) apply equally to the process started from some arbitrary time after corresponding changes in the stopping times.

We conclude this section with an implication of (2.1.3) and (2.1.4) that we often will use: (2.1.3) shows V_t cannot hit a level $a \in (-g/\gamma, V_0)$ before $(V_0 + g/\gamma)\, e^{-\gamma t} - g/\gamma$ does so at $\frac{1}{\gamma}\log\left(\frac{V_0+g/\gamma}{a+g/\gamma}\right)$. When $H_0 > 0$ and $\frac{1}{\gamma}\log\left(\frac{V_0+g/\gamma}{a+g/\gamma}\right) \leq \sigma(0)$, (2.1.4) shows V hits a exactly at $\frac{1}{\gamma}\log\left(\frac{V_0+g/\gamma}{a+g/\gamma}\right)$. In summary, for $V_0 > a$,

$$\tau_a^V \geq \frac{1}{\gamma}\log\left(\frac{V_0+g/\gamma}{a+g/\gamma}\right), \quad \tau_a^V = \frac{1}{\gamma}\log\left(\frac{V_0+g/\gamma}{a+g/\gamma}\right) \text{ for } H_0 > 0,\ \sigma(0) \geq \frac{1}{\gamma}\log\left(\frac{V_0+g/\gamma}{a+g/\gamma}\right). \tag{2.1.6}$$

We note this also implies $\left\{\tau_a^V \leq \sigma(0)\right\} = \left\{\tau_a^V \leq \sigma(0),\ \tau_a^V = \frac{1}{\gamma}\log\left(\frac{V_0+g/\gamma}{a+g/\gamma}\right)\right\}$ for $H_0 > 0, V_0 > a$.

2.1.3 Main results: Stability and renewal theory

At the heart of our proofs is the renewal structure of the process, which we now formalize. For any stopping time α, define $\tau_b^V(\alpha) = \inf\left\{t \geq \alpha \,\middle|\, V_t = b\right\}$, so that $\tau_b^V = \tau_b^V(0)$. Define a sequence of stopping times where the process (H_t, V_t) visits the point $(0, -g/(1+\gamma))$, which we call the renewal point, as follows: Set $a = -(g + g/2\gamma)/(1+\gamma)$ and $b = -(g - g/2(1+\gamma))/(1+\gamma)$. Define $\zeta_{-1} = 0$ and

$$\begin{aligned} \zeta &:= \zeta_0 = \inf\left\{t \geq \tau_a^V \wedge \tau_b^V \,\middle|\, (H_t, V_t) = (0, -g/(1+\gamma))\right\}, \\ \zeta_{n+1} &:= \inf\left\{t \geq \tau_a^V(\zeta_n) \wedge \tau_b^V(\zeta_n) \,\middle|\, (H_t, V_t) = (0, -g/(1+\gamma))\right\} \qquad n \geq 0. \end{aligned} \tag{2.1.7}$$

The choices of a and b are not too important as long as $a \in (-g/\gamma, -g/(1+\gamma))$ and $b \in (-g/(1+\gamma), 0)$, and their values are chosen as above for computational convenience.

We follow an approach similar to Banerjee et al. (2019) and Banerjee and Mukherjee (2019a) in showing that ζ (and hence each ζ_j) is integrable by decomposing the path between carefully chosen hitting times.

Theorem 2.1.1. *There exist constants* $t_0(\gamma, g), c > 0$ *such that*

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\zeta > t^2\right) \le e^{-c\,t}$$

for all $t > t_0(\gamma, g)$.

Since each ζ_j is integrable and the process is strong Markov, we have

$$\{\zeta_{j+1} - \zeta_j\}_{j\ge 0} \quad \text{and} \quad \left\{(H_t, V_t)_{t\in[\zeta_j,\zeta_{j+1})}\right\}_{j\ge 0} \qquad \text{are i.i.d.} \tag{2.1.8}$$

The existence and a representation of the stationary distribution, along with an ergodicity result for time averages, follows from the integrability of ζ and the observation (2.1.8) using the techniques developed in Ch. 10 of Thorisson (2000).

Theorem 2.1.2. *The solution to equation* (2.1.1) *has a unique stationary distribution* π*. For* $A \in \mathcal{B}(S)$,

$$\pi(A) = \frac{\mathbb{E}_{(0,-g/(1+\gamma))}\left(\int_0^{\zeta} \mathbb{1}_{(H_t,V_t)\in A}\, dt\right)}{\mathbb{E}_{(0,-g/(1+\gamma))}\ (\zeta)}.$$

In addition, for all $f : S \mapsto \mathbb{R}$ *bounded and measurable, we have*

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t P^s((h,\nu), f)\, ds = \int f(y)\, \pi(dy).$$

2.1.4 Main results: Path fluctuations, tail estimates and exponential ergodicity

The renewal structure laid out in Subsection 2.1.3 allows us to make statements about the long-time behavior of the system by studying its behavior in the (random) time interval $[0, \zeta)$, starting from the renewal point $(0, -g/(1+\gamma))$. An analysis of fluctuations in this random time interval translates to tail estimates for the stationary distribution π as displayed in Theorem 2.1.3. It also produces long-time oscillation estimates for V_t snd H_t, given in Theorem 2.1.4.

Theorem 2.1.3. *There exists positive constants $y'(\gamma, g), c', c$ such that for all $y > y'(\gamma, g)$,*

$$e^{-c'}\ e^{-4\,(1+\gamma)\left(y+\frac{g}{1+\gamma}\right)^2} \leq \pi\left(\mathbb{R}_+ \times (y, \infty)\right) \leq\ e^{c}\ e^{-\frac{1+\gamma}{8}\left(y+\frac{g}{1+\gamma}\right)^2}.$$

Also, there exists positive constants $x'(\gamma, g), C', C$ such that for all $x > x'(\gamma, g)$,

$$e^{-C'}e^{-\frac{4gx}{\gamma}} \leq \pi\left((x, \infty) \times \mathbb{R}\right) \leq e^{C}\ e^{-\frac{gx}{32\gamma}}.$$

Theorem 2.1.4. *For each $(h, \nu) \in S$, $\mathbb{P}_{(h,\nu)}$-a.s.,*

$$\begin{aligned} \frac{1}{\sqrt{2}\sqrt{1+\gamma}} \leq &\quad \limsup_{t\to\infty} \tfrac{V_t}{\sqrt{\log t}} \quad \leq 2\frac{1}{\sqrt{1+\gamma}} \\ \frac{\gamma}{2g} \leq &\quad \limsup_{t\to\infty} \tfrac{H_t}{\log t} \quad \leq 16\frac{\gamma}{g}. \end{aligned}$$

Remark 2.1.1. *Note that one cannot recover the analogous results in Banerjee et al. (2019) (see Theorems 2.1 and 2.2 in Banerjee et al. (2019)) by taking $\gamma \to 0$ in the results for the gap displayed in Theorems 2.1.3 and 2.1.4 above. This is because, when $\gamma = 0$, the primary contribution to the tail estimates for the stationary gap distribution comes from rare events when a large upward climb of the Brownian particle leads to a large positive value for the velocity of the inert particle. During such events the inert particle moves up rapidly, and the Brownian particle cannot 'keep up,' resulting in a large gap.*

In contrast when $\gamma > 0$, the viscosity term $-\gamma V$ ensures the inert particle moves more 'sluggishly' and cannot escape the Brownian particle. Thus, the large gaps arise when the Brownian particle escapes the inert particle by having a large downward fall. As we will show, the gap behaves like reflected Brownian motion with drift $-g/\gamma$ during such excursions.

The oscillation estimates obtained in Theorem 2.1.4 lead to a law of large numbers result for the trajectories S_t and X_t.

Theorem 2.1.5. *For each $(h, \nu) \in S$, $\mathbb{P}_{(h,\nu)}$-a.s.,*

$$\lim_{t\to\infty} \frac{S_t}{t} = \lim_{t\to\infty} \frac{X_t}{t} = -\frac{g}{1+\gamma}.$$

The next result shows that convergence to stationarity happens exponentially fast. Define the total variation norm $\|\cdot\|_{TV}$ for signed measures μ as $\|\mu\|_{TV} = \sup_{A\in\mathcal{B}(S)} |\mu|(A)$.

Theorem 2.1.6. *There exists a function $G : S \mapsto [1,\infty)$ and constants $D \in (0,\infty)$, $\lambda \in (0,1)$ such that for all $(h,\nu) \in S$ and $t \geq 0$,*

$$\|P^t\left((h,\nu),\cdot\,\right) - \pi\|_{TV} \leq G(h,\nu)\, D\lambda^t.$$

2.2 Existence and uniqueness of the process

We show the existence of a pathwise unique strong solution to equations (2.1.1) that is also a strong Markov process. The results proved here are for systems slightly more general than our current model. Consider the system of equations for $U = (U^1, U^2)$, with $U_0 = u = (u^1, u^2) \in \mathbb{R}_+ \times \mathbb{R}$ fixed,

$$\begin{aligned} dU_t^1 &= \varphi(U_t)\, dt - dB_t + dL_t^u \\ dU_t^2 &= \phi(U_t) dt + dL_t^u, \end{aligned} \tag{2.2.1}$$

where L_t^u is a continuous, non-negative, non-decreasing process such that $\int_0^t \mathbb{1}_{U_s^1=0} dL_s^u = L_t^u$ and $U_t^1 \geq 0$ for all t. A solution to (2.2.1) is therefore in $\mathcal{C}(\mathbb{R}_+, S)$, the space of continuous functions on $\mathbb{R}_+$ taking values in $S = \mathbb{R}_+ \times \mathbb{R}$. We write $\partial S = \{(u^1, u^2) \,:\, u^1 = 0\}$.

There has been substantial previous work in this direction. Ikeda and Watanabe (1989) establishes the existence of weak solutions and uniqueness in law for very general reflected systems. Existence and uniqueness results for inert drift systems have been addressed in Knight (2001) and White (2007) for models where the velocity is proportional to the local time and consequently is an increasing process. Bass et al. (2010) deals with weak existence and stationarity of inert drift systems on bounded domains. Our model differs qualitatively from those works, and we prove the existence of a pathwise unique strong solution to equations (2.1.1) that is also a strong Markov process. Under regularity conditions imposed on the drift (φ, ϕ), our proof adapts the standard procedure for existence and uniqueness of stochastic differential equations without reflection, with appropriate modifications to incorporate the reflection term.

We use the standard norm on $\mathcal{C}(\mathbb{R}_+, S)$, the continuous functions on the half-line taking values in $\mathbb{R}^d$, is given by $|f| = \sum_{n=1}^{\infty} \frac{1 \wedge |f|_n}{2^n}$, where $|\cdot|_n$ is the supremum norm on $\mathcal{C}([0,n], S)$. Also recall that a transition semigroup (Definition 6.1 in Le Gall (2013)) is a real-valued map $(t, u, f) \mapsto P^t(u, f)$ for $t \geq 0$, $u \in S$ and f bounded, measurable real-valued functions on S such that $A \mapsto P^t(u, \mathbb{1}_A)$ is a probability measure on the Borel σ-field of S for each t and u, $\left(P^t \circ P^s\right)(u, f) = P^{t+s}(u, f)$ for each $s, t \geq 0$, $P^0(u, f) = f(u)$ for all u and $(t, u) \mapsto P^t(u, f)$ is measurable for each such f.

Theorem 2.2.1. *Assume ϕ, φ are globally Lipschitz. Then for each initial condition $u = (u^1, u^2) \in S$ there exists a pathwise unique, strong solution to equations* (2.2.1) *in $\mathcal{C}(\mathbb{R}_+, S)$. The solution is a strong Markov process, and we have the Skorohod representation*

$$L_t^u = \sup_{\ell \leq t} \left(-u^1 + B_\ell - \int_0^\ell \varphi(U_s)\, ds \right) \vee 0.$$

Write $U(u)$ for the solution with $U_0 = u \in S$. The solution U may be chosen such that $u \mapsto U(u)$ is continuous in the topology of $\mathcal{C}(\mathbb{R}_+, S)$. $P^t(u, f) = \mathbb{E}_u\left(f\left(U_t\right)\right)$, where $f : S \mapsto \mathbb{R}$ is bounded and measurable, defines a transition semigroup.

Proof. We suppress the notation for initial conditions when the distinction is unnecessary. Define the stopping time and stopped processes

$$T_N = \inf\left\{t \geq 0 \,:\, |U_t| > N\right\}, \qquad U_t^N = U_{t \wedge T_N}, \qquad N \geq 1. \tag{2.2.2}$$

The drift vector and diffusion matrices of the stopped system satisfy the boundedness and Lipschitz conditions required in Theorem 7.2, Chapter IV.7 of Ikeda and Watanabe (1989), which shows that for each $N \geq 1$ there exists a weak solution U^N to (2.2.1) in $\mathcal{C}(\mathbb{R}_+, S)$, which has the strong Markov property and is unique in law. Skorohod's lemma (Lemma 6.14 Karatzas and Shreve (1991)) gives the representation for L^u stated in the theorem.

Uniqueness is in fact pathwise: Take $U^N, (U')^N$ to be two weak solutions, defining T'_N analogously to (2.2.2). We can assume that $U^N, (U')^N$ exist on a single filtered probability space and that the Brownian motions corresponding to the solutions are the same. Such a construction relies on regular conditional probabilities as described in Karatzas and Shreve (1991) Chapter 5D.

We use Gronwall's lemma in the typical way to show pathwise uniqueness. Set $\tilde{U}_t = U_{t\wedge T_N\wedge T'_N} - U'_{t\wedge T_N\wedge T'_N}$, suppressing the superscript N on $\tilde{U}$. Denote the local times corresponding to $U^N, (U')^N$ as L, L'. Fixing $T > 0$, we calculate for all $t \in [0,T]$

$$\begin{aligned}
&\mathbb{E}\left(|\tilde{U}_t|^2\right) \leq \mathbb{E}\left(\sup_{s\leq t}\left|\tilde{U}_s\right|^2\right) \leq K\left(\mathbb{E}\sup_{s\leq t}\left(L_{s\wedge T_N\wedge T'_N} - L'_{s\wedge T_N\wedge T'_N}\right)^2\right)\\
&+ K\,\mathbb{E}\left(\sup_{s\leq t}\left|\int_0^{s\wedge T_N\wedge T'_N}\varphi(U^N_\ell) - \varphi((U')^N_\ell)\,d\ell\right|^2 + \sup_{s\leq t}\left|\int_0^{s\wedge T_N\wedge T'_N}\phi(U^N_\ell) - \phi((U')^N_\ell)\,d\ell\right|^2\right)\\
&\leq K'\,\mathbb{E}\left(\sup_{s\leq t}\left|\int_0^{s\wedge T_N\wedge T'_N}\varphi(U^N_\ell) - \varphi((U')^N_\ell)\,d\ell\right|^2 + \sup_{s\leq t}\left|\int_0^{s\wedge T_N\wedge T'_N}\phi(U^N_\ell) - \phi((U')^N_\ell)\,d\ell\right|^2\right)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\leq K''T\,\mathbb{E}\left(\int_0^t |\tilde{U}_s|^2\,ds\right). \qquad (2.2.3)
\end{aligned}$$

$K, K', K'' > 0$ are constants not depending on T, u or N. The third inequality follows from the explicit form of L furnished by Skorohod's lemma, and the last from using Jensen's inequality and the Lipschitz property. By Gronwall's lemma, $\tilde{U}_t = 0$ on $[0,T]$. Since T was arbitrary, $U_{t\wedge T_N\wedge T'_N} = U'_{t\wedge T_N\wedge T'_N}$ for all t and each N. Calculations almost identical to (2.2.3) with U^N and $(U')^N$ in place of $\tilde{U}$, along with Gronwall's lemma again, imply $\lim_{N\to\infty} T_N = \infty$ and the same for T'_N. If $N_1 \leq N_2$, then $U^{N_1}_t = U^{N_2}_t$ for $t \leq T_{N_1}$, and we define a continuous process U for all time t such that $U_t = U^N_t$ for any N such that $t \leq T_N$. Similarly define U' for all time. Taking $N \to \infty$, we have $U_t = U'_t$ for all t, i.e. pathwise uniqueness holds.

Arguments based on regular conditional distributions used to prove the Yamada-Watanabe Theorem, and its Corollary 3.23, Chapter 5D of Karatzas and Shreve (1991), remain valid in the setting of Theorem 2.2.1 once weak existence and pathwise uniqueness are shown. Therefore U is the unique strong solution to (2.2.1).

To show $u \mapsto U(u)$ is continuous, we can follow almost exactly the argument as in the proof of Theorem 8.5 in Le Gall (2013). A minor change is needed because of the term L^u_t: consider the solutions $U(u)$ and $U(\widetilde{u})$ starting from distinct points u and $\widetilde{u}$. It suffices to show,

$$\mathbb{E}\left(\sup_{s\le t}\left|U^N(u)_s - U^N(\widetilde{u})_s\right|^2\right) \le K\left(|u-\widetilde{u}|^2 + T^2\int_0^t \mathbb{E}\left(\left|U^N(u)_s - U^N(\widetilde{u})_s\right|^2\right) ds\right) \quad (2.2.4)$$

for any $t \in [0,T]$, for all $T > 0$ and $N \ge 1$. Then, as in Theorem 8.5 of Le Gall (2013), continuity of $u \mapsto U(u)$ follows by sending $N \to \infty$ and using (2.2.4) to apply Gronwall's and Kolmogorov's continuity lemmas (Theorem 2.9 of Le Gall (2013)). However, (2.2.4) follows exactly as in the proof of (2.2.3) by replacing L, L' with $L^u, L^{\widetilde{u}}$ and $\tilde{U}$ with $U^N(u) - U^N(\widetilde{u})$ for any $u, \widetilde{u} \in S$. The strong Markov property for U holds since it does for U^N, by the calculation

$$\mathbb{1}_{T_N>\tau+t}\mathbb{E}\left(f\left(U_{\tau+t}(u)\right) \,\middle|\, \mathcal{F}_\tau\right) = \mathbb{1}_{T_N>\tau+t}\mathbb{E}\left(f\left(U_{\tau\wedge T_N+t}(u)\right) \,\middle|\, \mathcal{F}_{\tau\wedge T_N}\right) = \mathbb{1}_{T_N>\tau+t}P^t\left(U_{\tau\wedge T_N}, f\right),$$

for any finite stopping time τ and $t > 0$, and any bounded continuous f. As $N \to \infty$, $\mathbb{1}_{T_N>\tau+t} \to 1$ and the right-hand side tends to $P^t\left(U_\tau, f\right)$ by continuity. To check that $P^t(u, f)$ defines a transition semigroup, it remains only to show $t \mapsto P^t(u, f)$ is measurable for every u and bounded measurable f. However, the Dominated Convergence Theorem implies $t \mapsto P^t(u, f)$ is continuous for every bounded continuous f, and the proof is completed by a standard Monotone Class Theorem argument. □

2.3 Existence, uniqueness and representation of the stationary measure

This section is devoted to proving Theorems 2.1.1 and 2.1.2. We consider two cases. The case when the renewal point is approached from below, that is when the velocity hits the level $a < -g/(1+\gamma)$ before $b > -g/(1+\gamma)$, is considered in Subsection 2.3.1. Note that in this case, the first hitting time of the level $-g/(1+\gamma)$ by the velocity after hitting a corresponds to the renewal time, by (2.1.5). The other case when b is hit before a is considered in Subsection 2.3.2. In this case, the velocity can hit level $-g/(1+\gamma)$ after hitting b without the gap being zero at this hitting time. Thus, the velocity can fall below $-g/(1+\gamma)$ after hitting b, then approach $-g/(1+\gamma)$ from below for the renewal time to be attained. Integrability of the renewal time is proven by splitting the path into excursions between carefully chosen stopping times and estimating the duration of each such excursion.

2.3.1 Renewal point approached from below

Fix $a = -\left(g + g/2\gamma\right)/(1+\gamma) \in (-g/\gamma, -g/(1+\gamma))$, and $b = -\left(g - g/2(1+\gamma)\right)/(1+\gamma) \in (-g/(1+\gamma), 0)$. Define $\alpha_{-1} = 0$ and $\alpha_0 = \tau_a^V \wedge \tau_b^V$. If $\tau_b^V < \tau_a^V$, define $\alpha_j = \alpha_0$ for all $k \geq 0$ and $N^- = 0$. For $k \geq 0$, if $V_{\alpha_{3k}} = a$, define

$$\begin{aligned}
\alpha_{3k+1} &= \sigma(\alpha_{3k}) \\
\alpha_{3k+2} &= \inf\left\{t \geq \alpha_{3k+1} \,|\, V_t = -\left(g + g/4\gamma\right)/(1+\gamma)\right\} \\
\alpha_{3k+3} &= \inf\left\{t \geq \alpha_{3k+2} \,|\, V_t = -g/(1+\gamma) \quad \text{or} \quad a\right\} \\
N^- &= \inf\left\{k \geq 1 \,|\, V_{\alpha_{3k}} = -g/(1+\gamma)\right\}. \qquad (2.3.1)
\end{aligned}$$

If $V_{\alpha_{3k}} = -g/(1+\gamma)$, then $\alpha_j = \alpha_{3k}\ \forall\, j \geq 3k$. Unless otherwise stated, we assume $(H_0, V_0) = (0, -g/(1+\gamma))$. Lemma .0.3 says $\tau_a^V \wedge \tau_b^V < \infty$ a.s. with respect to the measure $\mathbb{P}_{(0,-g/(1+\gamma))}$. Lemma .0.2 shows $\alpha_1 < \infty$, and the proof of Lemma 2.3.2 applied successively to $\alpha_2 - \alpha_1$, $\alpha_3 - \alpha_2$ etc. will show the remaining α_j are finite as well.

Lemma 2.3.1. *For all $k \geq 0$,*

$$\mathbb{P}_{(0,-\frac{g}{1+\gamma})}\left(N^- > k\right) \leq \left(1 - \frac{1}{\sqrt{2\pi}} \frac{\frac{g}{\sqrt{2\gamma(1+\gamma)}}}{\left(\frac{g}{\sqrt{2\gamma(1+\gamma)}}\right)^2 + 1} \exp\left\{-\frac{g^2}{4\gamma(1+\gamma)^2}\right\}\right)^k.$$

Proof. Write $\epsilon = g/(2\gamma)$. For $(H_0, V_0) = \left(0, -\frac{g+\epsilon/2}{1+\gamma}\right)$ and $t < \tau^V_{-\frac{g+\epsilon}{1+\gamma}} \wedge \tau^V_{-\frac{g}{1+\gamma}}$, define

$$Y_t := -\frac{g+\epsilon/2}{1+\gamma} - t\frac{g}{1+\gamma} + \sup_{s \leq t}\left(B_s + s\frac{g}{1+\gamma}\right).$$

As $L_t \geq \sup_{s \leq t}\left(B_s + s\frac{g}{1+\gamma}\right)$ and $V_t \leq -g/(1+\gamma)$, therefore

$$V_t = V_0 - \gamma \int_0^t V_u du - gt + L_t \geq Y_t$$

for all $t < \tau^V_{-\frac{g+\epsilon}{1+\gamma}} \wedge \tau^V_{-\frac{g}{1+\gamma}}$. Note that Y_t cannot hit $-\frac{g+\epsilon}{1+\gamma}$ before time $\frac{\epsilon}{2g}$. If $\sup_{s\leq\epsilon/(4g)}\left(B_s + s\frac{g}{1+\gamma}\right) \geq \frac{3}{4}\frac{\epsilon}{1+\gamma}$, then Y_t and thus V_t must have hit $-\frac{g}{1+\gamma}$ by time $\epsilon/(4g)$. Thus,

$$\begin{aligned}
\mathbb{P}_{(0,-\frac{g}{1+\gamma})}\left(N^- > 1\right) &\leq \mathbb{P}_{\left(0,-\frac{g+\epsilon/2}{1+\gamma}\right)}\left(\tau^V_{-\frac{g+\epsilon}{1+\gamma}} < \tau^V_{-\frac{g}{1+\gamma}}\right) \\
&\leq \mathbb{P}\left(Y_t \text{ hits } -\frac{g+\epsilon}{1+\gamma} \text{ before } -\frac{g}{1+\gamma}\right) \leq \mathbb{P}\left(\sup_{s\leq\frac{\epsilon}{4g}}\left(B_s + s\frac{g}{1+\gamma}\right) < \frac{3}{4}\frac{\epsilon}{1+\gamma}\right) \\
&\leq \mathbb{P}\left(B_{\epsilon/4g} < \frac{1}{2}\frac{\epsilon}{1+\gamma}\right) \leq 1 - \frac{1}{\sqrt{2\pi}}\frac{\frac{\sqrt{\epsilon g}}{1+\gamma}}{\left(\frac{\sqrt{\epsilon g}}{1+\gamma}\right)^2 + 1} e^{-\frac{g\epsilon}{2(1+\gamma)^2}},
\end{aligned}$$

where the first inequality follows from the definition of N^- and the last inequality follows from a standard lower bound on the normal distribution function. Successive application of the strong Markov property yields the result. □

Lemma 2.3.2. *Fix $\gamma, g > 0$. There exist positive constants $t_0(\gamma, g), c$ such that for $j = 1, 2, 3$,$k \geq 0$*

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\alpha_{3k+j} - \alpha_{3k+j-1} > t\right) = \mathbb{P}_{(0,-g/(1+\gamma))}\left(\alpha_{3k+j} - \alpha_{3k+j-1} > t,\ N^- \geq k+1\right) \leq e^{-ct},$$

for all $t > t_0(\gamma, g)$.

Proof. By definition, if $\tau^V_a < \tau^V_b$ then $\alpha_0 = \tau^V_a$ and $\alpha_1 = \sigma(\tau^V_a)$. Lemma .0.4 shows there exists a positive constant c such that for t sufficiently large,

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\alpha_1 - \alpha_0 > t\right) = \mathbb{P}_{(0,-g/(1+\gamma))}\left(\sigma(\tau^V_a) - \tau^V_a > t,\ \tau^V_a < \tau^V_b\right) \leq e^{-ct}. \tag{2.3.2}$$

Now consider the difference $\alpha_{3k+1} - \alpha_{3k}$ for $k \geq 1$. If $N^- \leq k$ then $\alpha_{3k+1} - \alpha_{3k} = 0$. When $N^- \geq k+1$, α_{3k-1} is a point of increase of the velocity, since it has hit $-\frac{g+\frac{g}{4\gamma}}{1+\gamma}$ from below. By (2.1.5), $H_{\alpha_{3k-1}} = 0$. α_{3k} is the first time after α_{3k-1} that V hits a, and $\alpha_{3k+1} = \sigma(\alpha_k)$. In addition, $N^- \geq k+1$ implies the velocity starting from time α_{3k-1} does not hit $-g/(1+\gamma)$ before a and therefore does not hit $b > -g/(1+\gamma)$ before a either. Since $\left(H_{\alpha_{3k-1}}, V_{\alpha_{3k-1}}\right) \in \{0\} \times (a, b)$, we apply the strong Markov property at α_{3k-1} and Lemma .0.4 to show

$$\begin{aligned}\mathbb{P}_{(0,-g/(1+\gamma))}\left(\alpha_{3k+1}-\alpha_{3k}>t\right)&=\mathbb{P}_{(0,-g/(1+\gamma))}\left(\alpha_{3k+1}-\alpha_{3k}>t,\ N^{-}\geq k+1\right)\\&=\mathbb{E}_{(0,-g/(1+\gamma))}\left(\mathbb{1}_{N^{-}\geq k}\,\mathbb{P}_{\left(H_{\alpha_{3k-1}},V_{\alpha_{3k-1}}\right)}\left(\sigma(\tau_a^V)-\tau_a^V>t,\ \tau_a^V<\tau_{-g/(1+\gamma)}^V\right)\right)\\&\leq\sup_{\nu\in(a,b)}\mathbb{P}_{(0,\nu)}\left(\sigma(\tau_a^V)-\tau_a^V>t,\ \tau_a^V<\tau_b^V\right)\leq e^{-ct},\end{aligned}\tag{2.3.3}$$

for all t sufficiently large and some positive constant c. Using Lemma .0.6 with $\epsilon_0=\frac{g}{4\gamma}, h=0$ and applying the strong Markov property at α_{3k+1} gives for $k\geq 0$,

$$\begin{aligned}\mathbb{P}_{(0,-g/(1+\gamma))}\left(\alpha_{3k+2}-\alpha_{3k+1}>t\right)&=\mathbb{P}_{(0,-g/(1+\gamma))}\left(\alpha_{3k+2}-\alpha_{3k+1}>t,\ N^{-}\geq k+1\right)\\&\leq e^{C}\,e^{-t(g/\gamma)^2/32}.\end{aligned}\tag{2.3.4}$$

Apply Lemma .0.3 for the tail bound on the escape time of the interval $[a,-g/(1+\gamma)]$,

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\alpha_{3k+3}-\alpha_{3k+2}>t\right)\leq p^t,\tag{2.3.5}$$

holds for some $p\in(0,1)$ depending only on γ,g and for all t sufficiently large. Combining (2.3.2), (2.3.3), (2.3.4) and (2.3.5) completes the proof. □

Lemma 2.3.3. *There exist positive constants $t_0'(\gamma,g),c'$ such that*

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\zeta>t^2,\tau_a^V<\tau_b^V\right)=\mathbb{P}_{(0,-g/(1+\gamma))}\left(\alpha_{3N^-}>t^2,\tau_a^V<\tau_b^V\right)\leq e^{-c'\,t},$$

for all $t>t_0'(\gamma,g)$.

Proof. For any $k\geq 1$, if $N^-\leq k$, then $\alpha_{3(j+1)}=\alpha_{3j}$ for $j\geq k$. So for $t>0,n\geq 1$, the union bound shows

$$\begin{aligned}&\mathbb{P}_{(0,-g/(1+\gamma))}\left(1\leq N^-\leq n,\quad\sup_{1\leq k\leq(N^--1)}\left(\alpha_{3(k+1)}-\alpha_{3k}\right)>t\right)\\&\leq\sum_{k=0}^{n-1}\mathbb{P}_{(0,-g/(1+\gamma))}\left(\alpha_{3(k+1)}-\alpha_{3k}>t,\quad N^-\geq k+1\right),\end{aligned}\tag{2.3.6}$$

and

$$\begin{aligned}
\mathbb{P}_{(0,-g/(1+\gamma))}\left(\alpha_{3N^-} > t^2,\ \tau_a^V < \tau_b^V\right) &\le \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^- > t\right) \\
&\quad + \mathbb{P}_{(0,-g/(1+\gamma))}\left(\sum_{k=0}^{N^- -1}\left(\alpha_{3(k+1)} - \alpha_{3k}\right) > t^2,\ 0 < N^- \le t\right) \\
&\le \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^- > t\right) + \mathbb{P}_{(0,-g/(1+\gamma))}\left(1 \le N^- \le t,\ \sup_{1\le k\le(N^- -1)}\left(\alpha_{3(k+1)} - \alpha_{3k}\right) > t\right).
\end{aligned}$$

Using (2.3.6) and applying the bounds obtained in Lemma 2.3.1 and Lemma 2.3.2, we obtain

$$\begin{aligned}
\mathbb{P}_{(0,-g/(1+\gamma))}\left(\alpha_{3N^-} > t^2,\ \tau_a^V < \tau_b^V\right) &\le \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^- > t\right) \\
&\quad + \sum_{k=0}^{\lfloor t\rfloor -1} \mathbb{P}_{(0,-g/(1+\gamma))}\left(\alpha_{3(k+1)} - \alpha_{3k} > t,\ \ N^- \ge k+1\right) \\
&\le \left(1 - \frac{1}{\sqrt{2\pi}} \frac{\frac{g}{\sqrt{2\gamma}(1+\gamma)}}{\left(\frac{g}{\sqrt{2\gamma}(1+\gamma)}\right)^2 + 1}\ \exp\left\{-\frac{g^2}{4\gamma(1+\gamma)^2}\right\}\right)^{t-1} + te^C\, e^{-c\,(t-1)},
\end{aligned}$$

which gives the bound stated in the lemma for sufficiently large t. □

2.3.2 Renewal point approached from above

Our goal in this section is to bound $\mathbb{P}_{(0,-g/(1+\gamma))}\left(\zeta > t^2, \tau_b^V < \tau_a^V\right)$, the case where V rises to the level $b > \frac{-g}{1+\gamma}$ before the process returns to $(0, -g/(1+\gamma))$. Define $\beta_{-1} = 0$ and $\beta_0 = \tau_a^V \wedge \tau_b^V$. If $\tau_a^V < \tau_b^V$, define $\beta_j = \beta_0$ for all $k \ge 0$ and $N^+ = 0$. If $V_{\beta_{3k}} = b$ and $k \ge 0$,

$$\begin{aligned}
\beta_{3k+1} &= \inf\left\{t \ge \beta_{3k} \mid V_t = -\frac{g - \frac{g}{4(1+\gamma)}}{1+\gamma}\right\} \\
\beta_{3k+2} &= \sigma\left(\beta_{3k+1}\right) \wedge \inf\left\{t \ge \beta_{3k+1} \mid V_t = -\frac{g}{1+\gamma}\right\} \\
\beta_{3k+3} &= \inf\left\{t \ge \beta_{3k+2} \mid V_t = -\frac{g}{1+\gamma} \quad \text{or} \quad b\right\} \\
N^+ &= \inf\left\{k \ge 1 \mid V_{\beta_{3k}} = -\frac{g}{1+\gamma}\right\}.
\end{aligned} \tag{2.3.7}$$

If $V_{\beta_{3k}} = -\frac{g}{\gamma+1}$, then $\beta_j = \beta_{3k} \quad j \geq 3k$. The fact that almost surely, $\beta_j < \infty$ for all j and $N^+ < \infty$ can be shown using arguments similar to those succeeding (2.3.1).

For Section 2.3.1, in which $\tau_a^V < \tau_b^V$, we relied on the property (2.1.5) to imply that $H = 0$ at the time when V increases to $\frac{-g}{1+\gamma}$ from below. No such claim is possible when $\tau_b^V < \tau_a^V$ as V crosses $\frac{-g}{1+\gamma}$ from above. In this case, $\beta_{3N^+} \leq \zeta$, with equality only on the event $H_{\beta_{3N^+}} = 0$.

We therefore break this section into two parts: First, we derive bounds like those in Lemma 2.3.2 but for the differences $\{\beta_j - \beta_{j-1}\}_{j\geq 0}$ until the time β_{3N^+}, when V crosses $\frac{-g}{1+\gamma}$ from above. We then control the time taken after β_{3N^+} for the velocity to return to $\frac{-g}{1+\gamma}$ from below.

Lemma 2.3.4. *There exists a positive constants $t_0(\gamma, g), c$ such that*

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\beta_{3N^+} > t^2, \tau_b^V < \tau_a^V\right) \leq e^{-c\,t},$$

for all $t > t_0(\gamma, g)$.

Proof. Consider Lemma .0.5 with $u = -\frac{g - \frac{g}{4(1+\gamma)}}{1+\gamma}$ and $\nu = -\frac{g - \frac{g}{2(1+\gamma)}}{1+\gamma}$, so that $(1+\gamma)u + g = g/4(1+\gamma)$ and $u - \nu = -g/4(1+\gamma)^2$. The lemma shows there exist positive constants c and $t_0(\gamma, g)$ such that

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\beta_{3k+1} - \beta_{3k} > t, N^+ \geq k+1\right) \leq e^{-ct}, \tag{2.3.8}$$

for $t > t_0(\gamma, g)$. On $N^+ \geq k+1$, for $s \in [\beta_{3k+1}, \beta_{3k+2}]$, L_s is constant and the velocity decreases deterministically from $-\frac{g - \frac{g}{4(1+\gamma)}}{1+\gamma}$ at β_{3k+1} to some point bounded below by $-\frac{g}{1+\gamma}$ at time β_{3k+2}. This observation, along with (2.1.1), implies

$$-\frac{g}{1+\gamma} + \frac{g - \frac{g}{4(1+\gamma)}}{1+\gamma} \leq V_{\beta_{3k+2}} - V_{\beta_{3k+1}} \leq (\beta_{3k+2} - \beta_{3k+1})\left(-\frac{g}{1+\gamma}\right).$$

This implies

$$\beta_{3k+2} - \beta_{3k+1} \leq \frac{1}{4(1+\gamma)}. \tag{2.3.9}$$

For any $k \geq 0$, the strong Markov property at β_{3k+2} and Lemma .0.3 produce a constant $\widetilde{p} \in (0,1)$ depending only on γ, g such that

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\beta_{3k+3} - \beta_{3k+2} > t\right) \leq \widetilde{p}^{\,t}, \tag{2.3.10}$$

when t is sufficiently large. Arguments similar to those in Lemma 2.3.1 show for sufficiently large $n \geq 1$

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(N^{+} > n\right) \leq e^{-cn} \tag{2.3.11}$$

for a positive constant c. Using (2.3.8), (2.3.9), (2.3.10) and (2.3.11) in bounds along the lines of (2.3.6) and the subsequent equation with α_k replaced by β_k for each $k \geq 0$ and N^- replaced by N^+, the lemma follows. $\square$

We now consider the time needed for V to return to $-g/(1+\gamma)$ after β_{3N^+}. On $\tau_a^V < \tau_b^V$ define $\tilde{\alpha}_j = \tau_a^V$ for $j \geq -1$ and $N^\sharp = -1$. On $\tau_b^V < \tau_a^V$ define

$$\begin{aligned}
\tilde{\alpha}_{-1} &= \inf\{t \geq \beta_{3N^+} \mid H_t = 0 \quad \text{or} \quad V_t = a\} \\
\tilde{\alpha}_0 &= \inf\left\{t \geq \tilde{\alpha}_{-1} \mid V_t = -\frac{g}{1+\gamma} \quad \text{or} \quad V_t = a\right\}
\end{aligned}$$

If $V_{\tilde{\alpha}_0} = -g/(\gamma+1)$, define $\tilde{\alpha}_j = \tilde{\alpha}_0$ for all $j \geq 1$, otherwise define $\{\tilde{\alpha}_j\}_{j\geq 1}$ analogously to $\{\alpha_j\}_{j\geq 1}$ in (2.3.1), with $\tilde{\alpha}_1 = \sigma(\tilde{\alpha}_0)$ etc. Also define

$$N^\sharp = \inf\left\{k \geq 0 \,|\, V_{\tilde{\alpha}_{3k}} = -\frac{g}{\gamma+1}\right\},$$

from which it follows that $\zeta = \tilde{\alpha}_{3N^\sharp}$ when $\tau_b^V < \tau_a^V$.

Lemma 2.3.5. *There exist constants $t_0(\gamma, g), c > 0$ such that*

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\zeta - \beta_{3N^+} > t^2, \tau_b^V < \tau_a^V\right) = \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tilde{\alpha}_{3N^\sharp} - \beta_{3N^+} > t^2, \tau_b^V < \tau_a^V\right) \leq e^{-ct},$$

for $t > t_0(\gamma, g)$.

Proof. First note that if $\tau_b^V < \tau_a^V$ and $H_{\beta_{3N^+}} = 0$, then $\beta_{3N^+} = \tilde{\alpha}_{-1} = \tilde{\alpha}_0$ and $N^\sharp = 0$. Since $V_{\beta_{3N^+}} = -g/(1+\gamma)$, in this case we have $\tilde{\alpha}_{3N^\sharp} = \beta_{3N^\sharp} = \beta_{3N^+} = \zeta$.

Henceforth we consider the case $\tau_b^V < \tau_a^V, H_{\beta_{3N^+}} > 0$. Arguing as in Lemma 2.3.1, we have for $n \geq 0$ and a positive constant c,

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(N^\sharp > n\right) \leq e^{-cn}. \tag{2.3.12}$$

Since L_s is constant on $[\beta_{3N^+}, \tilde{\alpha}_{-1}]$, we proceed as in (2.3.9) to show there exists a constant $C > 0$ such that

$$\tilde{\alpha}_{-1} - \beta_{3N^+} \leq C. \tag{2.3.13}$$

In the following analysis, there are two cases to consider: $H_{\tilde{\alpha}_{-1}} = 0$ and $H_{\tilde{\alpha}_{-1}} > 0$. In the former situation, $V_{\tilde{\alpha}_{-1}} \in (a, -g/(1+\gamma))$, so we use Lemma .0.3 and the strong Markov property at $\tilde{\alpha}_{-1}$ to show there exists $\bar{p} \in (0,1)$ depending on γ, g such that for t sufficiently large,

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tilde{\alpha}_0 - \tilde{\alpha}_{-1} > t,\ H_{\tilde{\alpha}_{-1}} = 0\right) \leq \sup_{\nu \in (a,-g/(1+\gamma))} \mathbb{P}_{(0,\nu)}\left(t < \tau_a^V \wedge \tau_{-g/(1+\gamma)}^V\right) \leq \bar{p}^t. \tag{2.3.14}$$

The analysis of $\{\tilde{\alpha}_{j+1} - \tilde{\alpha}_j\}_{j \geq 0}$ can be done exactly as that of $\{\alpha_{j+1} - \alpha_j\}_{j \geq 3}$ performed in Lemma 2.3.2 using Lemmas .0.3, .0.4 and .0.6. Thus we proceed as in Lemma 2.3.3, using $N^\sharp$ in place of N^- and (2.3.12) instead of Lemma 2.3.1, to obtain a positive constant c such that for t sufficiently large

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tilde{\alpha}_{3N^\sharp} - \beta_{3N^+} > t^2, \tau_b^V < \tau_a^V,\ H_{\tilde{\alpha}_{-1}} = 0\right) \leq e^{-ct}. \tag{2.3.15}$$

Now consider $H_{\tilde{\alpha}_{-1}} > 0$. In that case, $\tilde{\alpha}_{-1} = \tilde{\alpha}_0 = \inf\{s \geq \beta_{3N^+} \mid V_s = a\}$. The methods of Lemma 2.3.2 fail to bound the probability of $\alpha_1 - \alpha_0 > t$ since in principle $H_{\tilde{\alpha}_{-1}}$ might be quite large. We first apply the union bound to show for $k \geq 1$,

$$\begin{aligned}
&\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tilde{\alpha}_1 - \tilde{\alpha}_0 > t,\ H_{\tilde{\alpha}_{-1}} > 0,\ N^+ = k\right) \\
&\qquad \leq \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^+ \geq k,\ \beta_{3k-2} - \beta_{3k-3} > t\right) \\
&\qquad + \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tilde{\alpha}_1 - \tilde{\alpha}_0 > t,\ H_{\tilde{\alpha}_{-1}} > 0,\ N^+ = k,\ \beta_{3k-2} - \beta_{3k-3} \leq t\right) \\
&\qquad \leq \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^+ \geq k,\ \beta_{3k-2} - \beta_{3k-3} > t\right) \\
&\qquad + \mathbb{P}_{(0,-g/(1+\gamma))}\big(\tilde{\alpha}_1 - \tilde{\alpha}_0 > t,\ H_{\tilde{\alpha}_{-1}} > 0, \\
&\qquad N^+ = k,\ \beta_{3k-2} - \beta_{3k-3} \leq t,\ H_{\beta_{3k-2}} \leq tg/4\gamma(1+\gamma)\big) \\
&\qquad + \mathbb{P}_{(0,-g/(1+\gamma))}\big(\tilde{\alpha}_1 - \tilde{\alpha}_0 > t,\ H_{\tilde{\alpha}_{-1}} > 0, \\
&\qquad N^+ = k,\ \beta_{3k-2} - \beta_{3k-3} \leq t,\ H_{\beta_{3k-2}} > tg/4\gamma(1+\gamma)\big).
\end{aligned} \tag{2.3.16}$$

Recall that β_{3k-3} is a point of increase of V to the level $b = -\left(g - g/2(1+\gamma)\right)/(1+\gamma)$, so (2.1.5) shows $H_{\beta_{3k-3}} = 0$. β_{3k-2} is the first time after β_{3k-3} the velocity falls to $-\left(g - g/4(1+\gamma)\right)/(1+\gamma)$. Using Lemma .0.5 and the strong Markov property at β_{3k-3}, there exists constants $c, t_0(\gamma, g) > 0$ such that for all $t > t_0(\gamma, g)$,

$$\begin{aligned}
&\mathbb{P}_{(0,-g/(1+\gamma))}\left(N^+ \geq k,\ \beta_{3k-2} - \beta_{3k-3} > t\right) \\
&\qquad = \mathbb{E}_{(0,-g/(1+\gamma))}\left(N^+ \geq k,\ \mathbb{P}_{(H_{\beta_{3k}}, V_{\beta_{3k}})}\left(\tau^V_{-(g-g/4(1+\gamma))/(1+\gamma)} > t\right)\right) \\
&\qquad \leq \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^+ \geq k\right) e^{-ct}. \qquad (2.3.17)
\end{aligned}$$

When $N^+ = k$ and $H_{\beta_{3N^+}}, H_{\tilde{\alpha}_{-1}} > 0$, the velocity starting from β_{3k-3} makes an excursion from b to $-\frac{g-g/4(1+\gamma)}{1+\gamma}$ at time β_{3k-2}, then to a without returning to b. Therefore,

$$\tilde{\alpha}_0 = \tilde{\alpha}_{-1} = \inf\{s \geq \beta_{3k-2} \mid V_s = a\} < \inf\{s \geq \beta_{3k-2} \mid V_s = b\}, \qquad \text{when } N^+ = k. \qquad (2.3.18)$$

Using (2.3.18), the strong Markov property at β_{3k-2} and Lemma .0.4 we obtain $t_1(\gamma, g) > 0$, which we take to be larger than $t_0(\gamma, g)$, such that for $t > t_1(\gamma, g)$,

$$\begin{aligned}
&\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tilde{\alpha}_1 - \tilde{\alpha}_0 > t,\ H_{\tilde{\alpha}_{-1}} > 0,\ N^+ = k, \beta_{3k-2} - \beta_{3k-3} \leq t,\ H_{\beta_{3k-2}} \leq tg/4\gamma(1+\gamma)\right) \\
&\qquad \leq \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^+ \geq k\right) \sup_{(h,\nu)\in[0,tg/4\gamma(1+\gamma)]\times(a,b)} \mathbb{P}_{(h,\nu)}\left(\sigma(\tau^V_a) - \tau^V_a > t,\ \tau^V_a < \tau^V_b\right) \\
&\qquad \leq \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^+ \geq k\right) e^{-ct}. \qquad (2.3.19)
\end{aligned}$$

Fix $t > t_1(\gamma, g)$ and set $(H_0, V_0) = (0, b)$. System equations (2.1.1) show $H_u + V_u/\gamma = b/\gamma - ug/\gamma - B_u + (1 + 1/\gamma)L_u$. When $u < t \wedge \tau^V_{-(g-g/4(1+\gamma))/(1+\gamma)}$, we have $S_u \geq -u\left(g - g/4(1+\gamma)\right)/(1+\gamma)$ and $L_u \leq \sup_{s\leq u}\left(B_s + s\left(g - g/4(1+\gamma)\right)/(1+\gamma)\right) \leq \sup_{s\leq u} B_s +$

$u\left(g-g/4(1+\gamma)\right)/(1+\gamma)$. As a result, with $c' = b/\gamma + \left(g-g/4(1+\gamma)\right)/\gamma(1+\gamma) > 0$,

$$
\begin{aligned}
H_u = H_u + V_u/\gamma - V_u/\gamma &\le c' - B_u - ug/\gamma + (1+1/\gamma)L_u \\
&\le c' - B_u + u\left((1+1/\gamma)\left(g-g/4(1+\gamma)\right)/(1+\gamma) - g/\gamma\right) + (1+1/\gamma)\sup_{s\le u} B_s \\
= c' - B_u - ug/4\gamma(1+\gamma) + (1+1/\gamma)\sup_{s\le u} B_s &\le c' + \sup_{s\le t}(-B_s) + (1+1/\gamma)\sup_{s\le t} B_s. \quad (2.3.20)
\end{aligned}
$$

From (2.3.20), we conclude that if $\tau^V_{-(g-g/4(1+\gamma))/(1+\gamma)} \le t$, then $H_{\tau^V_{-(g-g/4(1+\gamma))/(1+\gamma)}} > tg/4\gamma(1+\gamma)$ implies $\sup_{s\le t}(-B_s) + (1+1/\gamma)\sup_{s\le t} B_s > tg/4\gamma(1+\gamma) - c'$. We choose $t_1(\gamma,g)$ large enough that $t_1(\gamma,g) - c' > 0$. Now the strong Markov property at β_{3k-3}, (2.3.20) and Gaussian tail bounds show there exists a $t_2(\gamma,g) > t_1(\gamma,g)$ and constants $C, C', c, c' > 0$ such that that for $t > t_2(\gamma,g)$,

$$
\begin{aligned}
&\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tilde\alpha_1 - \tilde\alpha_0 > t,\ H_{\tilde\alpha_{-1}} > 0,\ N^+ = k, \beta_{3k-2} - \beta_{3k-3} \le t,\ H_{\beta_{3k-2}} > tg/4\gamma(1+\gamma)\right) \\
&\quad \le \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^+ \ge k\right)\ \mathbb{P}_{(0,b)}\left(H_{\tau^V_{-\frac{g-g/4(1+\gamma)}{1+\gamma}}} > tg/4\gamma(1+\gamma),\ \tau^V_{-(g-g/4(1+\gamma))/(1+\gamma)} \le t\right) \\
&\quad \le \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^+ \ge k\right)\ \Bigg[\mathbb{P}\left(\sup_{s\le t}(-B_s) > tg/8\gamma(1+\gamma) - c'/2\right) \\
&\qquad\qquad + \mathbb{P}\left(\sup_{s\le t}(B_s) > \frac{\gamma}{1+\gamma}\left(tg/8\gamma(1+\gamma) - c'/2\right)\right)\Bigg] \\
&\qquad\qquad\qquad \le \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^+ \ge k\right)\ \left[e^{C'}\, e^{-c't} + e^{C}\, e^{-ct}\right]. \quad (2.3.21)
\end{aligned}
$$

By (2.3.11), $\mathbb{E}_{(0,-g/(1+\gamma))}\, N^+ \le e^C$ for a positive constant C. We apply (2.3.17), (2.3.19) and (2.3.21) to (2.3.16) and sum over k to obtain positive constants c, $t_2(\gamma,g)$ such that,

$$
\begin{aligned}
&\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tilde\alpha_1 - \tilde\alpha_0 > t,\ H_{\tilde\alpha_{-1}} > 0\right) \\
&\qquad = \sum_{k=1}^{\infty} \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tilde\alpha_1 - \tilde\alpha_0 > t,\ H_{\tilde\alpha_{-1}} > 0,\ N^+ = k,\ N^+ \ge k\right) \\
&\qquad\qquad \le e^{-ct} \sum_{k=1}^{\infty} \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^+ \ge k\right) \le e^C\, e^{-ct}, \quad (2.3.22)
\end{aligned}
$$

for $t > t_2(\gamma,g)$. Now the analysis of $\{\tilde\alpha_{j+1} - \tilde\alpha_j\}_{j\ge1}$ can be done exactly as that of $\{\alpha_{j+1} - \alpha_j\}_{j\ge1}$ performed in Lemma 2.3.2. Once again we argue as in Lemma 2.3.3, using $N^\sharp$ in place of N^-

and (2.3.12) instead of Lemma 2.3.1, to obtain a positive constant c such that for t sufficiently large

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tilde{\alpha}_{3N^\sharp} - \beta_{3N^+} > t^2, \tau_b^V < \tau_a^V,\ H_{\tilde{\alpha}_{-1}} > 0\right) \leq e^{-ct}. \tag{2.3.23}$$

(2.3.15) and (2.3.23) prove the lemma. □

Now, we have all the tools needed to prove Theorems 2.1.1 and 2.1.2.

Proof of Theorem 2.1.1. This is a direct consequence of Lemmas 2.3.3, 2.3.4 and 2.3.5. □

Proof of Theorem 2.1.2. Theorem 2.1.1 and the strong Markov property, proven to hold in Theorem 2.2.1, show that the system is 'classical regenerative' and satisfies the conditions of Theorem 2.1, Chapter 10, in Thorisson (2000), which gives a stationary measure of the stated form.

Now we prove the second claim of the theorem, which also implies uniqueness of the stationary measure. Define $N_t = \sup\{k \geq 0 : \zeta_k \leq t\}$, the number of renewals before time t. It is enough to show the claim for bounded, non-negative f. Recalling $\zeta = \zeta_0$,

$$\begin{aligned}\int_0^{\zeta\wedge t} f(H_s, V_s)\, ds + \mathbb{1}_{\zeta\leq t}\sum_{k=1}^{N_t}\int_{\zeta_{k-1}}^{\zeta_k} f(H_s,V_s)\, ds &\leq \int_0^t f(H_s,V_s)\, ds \\ &\leq \int_0^{\zeta} f(H_s,V_s)\, ds + \sum_{k=1}^{N_t+1}\int_{\zeta_{k-1}}^{\zeta_k} f(H_s,V_s)\, ds. \end{aligned} \tag{2.3.24}$$

Now the Strong Law of Large Numbers implies $N_t/t \to 1/\mathbb{E}_{(0,-g/(1+\gamma))}(\zeta) > 0$ (e.g. Theorem 2.4.6 of Durrett (2010)). Using (2.1.8) and applying the Law of Large Numbers in (2.3.24) completes the proof. □

2.4 Fluctuation bounds on a renewal interval

We show that on the interval $[0, \zeta]$, the probability of the velocity hitting a large value y has Gaussian decay in y and the corresponding probability of the gap hitting a large value x has exponential decay in x. These estimates directly imply bounds on the tails of the stationary measure given in Theorem 2.1.3 via the representation in Theorem 2.1.2, and also produce the oscillation estimates stated in Theorem 2.1.4.

2.4.1 Velocity bounds

Lemma 2.4.1. *There exists a positive constant c such that*

$$\mathbb{P}_{(0,y)}\left(\tau_{2y}^V < \tau_{y/2}^V\right) \le e^c e^{-(1+\gamma)(y+\frac{g}{1+\gamma})^2},$$

for all $y > 0$.

Proof. Define $F_t = y + \sup_{u\le t}(B_u - uy/2) - (g + y\gamma/2)t$. For $t < \tau_{y/2}^V \wedge \tau_{2y}^V$, the system equations (2.1.1) and the Skorohod representation (2.1.2) imply $V_t \le F_t$. Hence,

$$\begin{aligned}\mathbb{P}_{(0,y)}\left(\tau_{2y}^V < \tau_{y/2}^V\right) \le \mathbb{P}\left(F_t \text{ hits 2y before y/2}\right) \le \mathbb{P}\left(\sup_{t<\infty}(B_t - ty/2 - (g + y\gamma/2)t) \ge y\right)\\ = \exp\left\{-2y(g + (1+\gamma)y/2)\right\} = e^{\frac{g^2}{1+\gamma}} e^{-(1+\gamma)(y+\frac{g}{1+\gamma})^2}.\end{aligned}$$

The second inequality comes from the fact that the time at which F_t hits $2y$ must be a point of increase of $\sup_{u\le t}(B_u - uy/2)$. The first equality uses the fact that for any $u > 0$, $\sup\limits_{s<\infty}(B_s - su) \overset{d}{=} \text{Exponential}(2u)$ (see Chapter 3.5 of Karatzas and Shreve (1991)). □

Fix any $y > 0$ and choose the starting configuration $(H_0, V_0) = (0, y)$. Define the following sequence of stopping times: $\tau_0 = 0$ and for $k \ge 0$,

$$\begin{aligned}\tau_{2k+1} &= \inf\{t \ge \tau_{2k} \,|\, V_t = 2y \quad \text{or} \quad y/2\} \qquad \text{if } V_{\tau_{2k}} \ne -g/(1+\gamma), 2y\\ &= \tau_{2k} \qquad \text{otherwise}\\ \tau_{2k+2} &= \inf\{t \ge \tau_{2k} \,|\, V_t = y \quad \text{or} \quad -g/(1+\gamma)\} \qquad \text{if } V_{\tau_{2k+1}} \ne -g/(1+\gamma), 2y\\ &= \tau_{2k+1} \qquad \text{otherwise}\\ J_y &= \min\{k \ge 1 \,|\, V_{\tau_{2k}} = -g/(1+\gamma) \quad \text{or} \quad 2y\}.\end{aligned} \tag{2.4.1}$$

Lemma 2.4.2. *There exist positive constants $y'(\gamma, g)$ and $p(\gamma, g) \in (0,1)$ such that*

$$\mathbb{P}_{(0,y)}\left(J_y > n\right) \le p(\gamma, g)^n,$$

for $y > y'(\gamma, g)$ and $n \ge 0$.

Proof. For any $y > g/(1+\gamma)$, applying the strong Markov property at τ_1, observe that

$$\begin{aligned}\mathbb{P}_{(0,y)}\left(J_y > 1\right) = \mathbb{P}_{(0,y)}\left(V_{\tau_1} = y/2,\ V_{\tau_2} = y\right) \le \mathbb{P}_{(0,y/2)}\left(\tau^V_y < \tau^V_{-\frac{g}{1+\gamma}}\right) \\ \le \mathbb{P}_{(0,y/2)}\left(\tau^V_y < \tau^V_{\frac{g}{1+\gamma}}\right) + \mathbb{P}_{(0,y)}\left(\tau^V_{\frac{g}{1+\gamma}} < \tau^V_y < \tau^V_{-\frac{g}{1+\gamma}}\right). \end{aligned} \quad (2.4.2)$$

We bound the second probability on the right hand side by applying the strong Markov property at the stopping time $\inf\left\{t \ge \sigma\left(\tau^V_{\frac{g}{1+\gamma}}\right) \,\middle|\, V_t = g/(1+\gamma)\right\}$ to obtain

$$\mathbb{P}_{(0,y)}\left(\tau^V_{\frac{g}{1+\gamma}} < \tau^V_y < \tau^V_{-\frac{g}{1+\gamma}}\right) \le \mathbb{P}_{(0,\frac{g}{1+\gamma})}\left(\tau^V_y < \tau^V_{-\frac{g}{1+\gamma}}\right).$$

Using this in (2.4.2), we obtain

$$\mathbb{P}_{(0,y)}\left(J_y > 1\right) \le \mathbb{P}_{(0,y/2)}\left(\tau^V_y < \tau^V_{\frac{g}{1+\gamma}}\right) + \mathbb{P}_{(0,\frac{g}{1+\gamma})}\left(\tau^V_y < \tau^V_{-\frac{g}{1+\gamma}}\right). \quad (2.4.3)$$

To estimate the first probability on the right hand side of (2.4.3), note that if $V_0 = y/2$ and $t < \tau^V_y \wedge \tau^V_{\frac{g}{1+\gamma}}$,

$$V_t \le y/2 - t\left(\gamma\frac{g}{1+\gamma} + g\right) + \sup_{s \le t}\left(B_s - s\frac{g}{1+\gamma}\right) := y/2 + Z_t.$$

Arguing as in Lemma 2.4.1,

$$\begin{aligned}\mathbb{P}_{(0,y/2)}\left(\tau^V_y < \tau^V_{\frac{g}{1+\gamma}}\right) \le \mathbb{P}\left(y/2 + Z_t \quad \text{hits } y \text{ before } \frac{g}{1+\gamma}\right) \\ \le \mathbb{P}\left(\sup_{t<\infty}\left(B_t - 2gt\right) > y/2\right) = e^{-2gy}. \end{aligned} \quad (2.4.4)$$

To estimate the second probability in (2.4.3), observe that for any $y \ge 2g/(1+\gamma)$,

$$\begin{aligned}&\mathbb{P}_{(0,\frac{g}{1+\gamma})}\left(\tau^V_{-\frac{g}{1+\gamma}} < \tau^V_y\right) \\ &\ge \mathbb{P}_{(0,\frac{g}{1+\gamma})}\left(\tau^V_{-\frac{g}{1+\gamma}} < \tau^V_{\frac{2g}{1+\gamma}}\right) \ge \mathbb{P}_{(0,\frac{g}{1+\gamma})}\left(\tau^H_1 < \tau^V_{\frac{2g}{1+\gamma}},\ \tau^V_{(-g/\gamma,-g/(1+\gamma)]} \in \left[\tau^H_1, \tau^V_{\frac{2g}{1+\gamma}}\right]\right) \\ &\ge \mathbb{P}_{(0,\frac{g}{1+\gamma})}\left(\tau^H_1 < \tau^V_{\frac{2g}{1+\gamma}}\right) \times \inf_{v \in (-g/(1+\gamma), 2g/(1+\gamma))} \mathbb{P}_{(1,v)}\left(\tau^V_{-\frac{g}{1+\gamma}} < \tau^V_{\frac{2g}{1+\gamma}}\right) \end{aligned} \quad (2.4.5)$$

where the last inequality follows by applying the strong Markov property at τ_1^H. Recalling that for any $t \geq 0$

$$H_t \geq H_0 - \frac{gt}{\gamma} - B_t + L_t \geq H_0 - \frac{gt}{\gamma} - B_t, \tag{2.4.6}$$

and for $V_0 = g/(1+\gamma)$, as $V_t > -g/\gamma$ for all t,

$$V_t = V_0 - \int_0^t (\gamma V_s + g)\, ds + L_t \leq \frac{g}{1+\gamma} + \sup_{u \leq t} \left(-H_0 + B_u + \frac{gu}{\gamma} \right),$$

we obtain

$$\mathbb{P}_{(0,\frac{g}{1+\gamma})} \left(\tau_1^H < \tau^V_{\frac{2g}{1+\gamma}} \right) \geq \mathbb{P} \left(B_1 \leq -1 - g/\gamma,\ \sup_{u \leq 1} \left(B_u + \frac{gu}{\gamma} \right) < \frac{g}{1+\gamma} \right) = p_1(\gamma, g) > 0. \tag{2.4.7}$$

Now we bound the second term in the product on the right hand side of (2.4.5) from below. Choosing starting point $(H_0, V_0) = (1, v)$ for any $v \in (-g/(1+\gamma), 2g/(1+\gamma))$, recall from (2.1.4) that for any $t > 0$, if $\sigma(0) > t$, then

$$V_t \leq \left(\frac{2g}{1+\gamma} + g/\gamma \right) e^{-\gamma t} - g/\gamma.$$

The right side equals $-g/(1+\gamma)$ when $t = \gamma^{-1} \log(1+3\gamma)$ and hence, if $\sigma(0) \geq \gamma^{-1} \log(1+3\gamma)$, then $\tau^V_{-\frac{g}{1+\gamma}} \leq \gamma^{-1} \log(1+3\gamma)$. Thus, from (2.4.6), if $B_t \leq \frac{1}{2} - \frac{gt}{\gamma}$ for all $t \leq \gamma^{-1} \log(1+3\gamma)$, then $\sigma(0) \geq \gamma^{-1} \log(1+3\gamma)$ and thus $\sigma(0) \geq \tau^V_{-\frac{g}{1+\gamma}}$. As the velocity is strictly decreasing on the interval $[0, \sigma(0)]$, for any $v \in (-g/(1+\gamma), 2g/(1+\gamma))$,

$$\begin{aligned} &\mathbb{P}_{(1,v)} \left(\tau^V_{-\frac{g}{1+\gamma}} < \tau^V_{\frac{2g}{1+\gamma}} \right) \\ &\qquad \geq \mathbb{P}_{(1,v)} \left(\tau^V_{-\frac{g}{1+\gamma}} \leq \sigma(0) \right) \geq \mathbb{P} \left(\sup_{u \leq \gamma^{-1} \log(1+3\gamma)} \left(B_u + \frac{gu}{\gamma} \right) \leq \frac{1}{2} \right) = p_2(\gamma, g) > 0. \end{aligned} \tag{2.4.8}$$

Using (2.4.7) and (2.4.8) in (2.4.5), we obtain

$$\mathbb{P}_{(0,\frac{g}{1+\gamma})} \left(\tau^V_{-\frac{g}{1+\gamma}} < \tau^V_y \right) \geq p_1(\gamma, g) p_2(\gamma, g) > 0 \tag{2.4.9}$$

Using (2.4.4) and (2.4.9) in (2.4.3)

$$\mathbb{P}_{(0,y)}\left(J_y > 1\right) \le e^{-2gy} + (1 - p_1(\gamma, g)p_2(\gamma, g)). \tag{2.4.10}$$

Choosing any $p(\gamma, g) \in ((1 - p_1(\gamma, g)p_2(\gamma, g)), 1)$ and $y'(\gamma, g) \ge 2g/(1+\gamma)$ such that the right hand side of (2.4.10) is bounded above by $p(\gamma, g)$ for all $y \ge y'(\gamma, g)$, the assertion of the lemma follows for $n = 1$. The result for $n \ge 2$ follows by induction upon using the strong Markov property at τ_{2n-1}. □

Theorem 2.4.3. *There exist positive constants $y'(\gamma, g), c, C$ such that*

$$e^{-C}\, e^{-2(1+\gamma)(y+g/(1+\gamma))^2} \le \mathbb{P}_{(0,-g/(1+\gamma))}\left(\sup_{[0,\zeta]} V_t \ge y\right) \le e^{c} e^{-\frac{1+\gamma}{4}(y+g/(1+\gamma))^2},$$

for all $y > y'(\gamma, g)$.

Proof. Consider $\{\tau_i\}_{i\ge1}$ and J_y defined in (2.4.1). Then, choosing $y'(\gamma, g)$ to be the same constant as in Lemma 2.4.2, for any $y > y'(\gamma, g)$, we apply the strong Markov property at stopping time τ^V_y to obtain

$$\begin{aligned}\mathbb{P}_{(0,-g/(1+\gamma))}\left(\sup_{[0,\zeta]} V_t \ge 2y\right) &\le \mathbb{P}_{(0,y)}\left(\tau^V_{2y} < \tau^V_{-g/(1+\gamma)}\right)\\ &\le \sum_{k=0}^{\infty} \mathbb{P}_{(0,y)}\left(\sup_{[\tau_{2k},\tau_{2k+1}]} V_t \ge 2y,\ J_y > k\right).\end{aligned} \tag{2.4.11}$$

By applying the strong Markov property at time τ_{2k},

$$\mathbb{P}_{(0,y)}\left(\sup_{[\tau_{2k},\tau_{2k+1}]} V_t \ge 2y,\ J_y > k\right) = \mathbb{P}_{(0,y)}\left(\tau^V_{2y} < \tau^V_{y/2}\right)\mathbb{P}_{(0,y)}\left(J_y > k\right).$$

Using this in (2.4.11), we obtain

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\sup_{[0,\zeta]} V_t \ge 2y\right) \le \mathbb{P}_{(0,y)}\left(\tau^V_{2y} < \tau^V_{y/2}\right)\ \mathbb{E}_{(0,-g/(1+\gamma))}\ (J_y)$$

from which the upper bound claimed in the theorem (with $2y$ in place of y) follows for $y > y'(\gamma, g)$ from Lemmas 2.4.1 and 2.4.2.

For the lower bound, we first show for $b = -\frac{g - \frac{g}{2(1+\gamma)}}{1+\gamma}$, as in the definition of ζ, that for y sufficiently large,

$$\mathbb{P}_{(0,b)}\left(\tau_y^V < \tau_{-\frac{g}{1+\gamma}}^V\right) \geq e^{-2(1+\gamma)\left(y+\frac{g}{1+\gamma}\right)^2}. \tag{2.4.12}$$

Choosing the starting point $(H_0, V_0) = (0, b)$, for $t < \tau_{-\frac{g}{1+\gamma}}^V \wedge \tau_y^V$, we have

$$V_t = b - \int_0^t (\gamma V_s + g)\, ds + L_t \geq b - (\gamma y + g)t + \sup_{u \leq t}(B_u - yu) \geq b - t\left((1+\gamma)\, y + g\right) + B_t := F_t.$$

We also note that if $s(v) = e^{2v((1+\gamma)y+g)}$, then $s(F_t)$ is a bounded martingale on $t < \tau_{-\frac{g}{1+\gamma}}^V \wedge \tau_y^V$. By the optional stopping theorem,

$$\begin{aligned}\mathbb{P}_{(0,b)}\left(\tau_y^V < \tau_{-\frac{g}{1+\gamma}}^V\right) &\geq \mathbb{P}_{(0,b)}\left(F_t \text{ hits } y \text{ before } -\frac{g}{1+\gamma}\right)\\ &= \frac{s(b) - s\left(-\frac{g}{1+\gamma}\right)}{s(y) - s\left(-\frac{g}{1+\gamma}\right)} \geq e^{-2(1+\gamma)\left(y+\frac{g}{1+\gamma}\right)^2}\end{aligned} \tag{2.4.13}$$

for y sufficiently large. This proves (2.4.12). Recall $a = -\frac{g+\frac{g}{2\gamma}}{1+\gamma} < -\frac{g}{1+\gamma}$ in the definition of ζ. Note that if $\tau_a^V < \tau_b^V$, then the renewal time ζ corresponds to the first hitting time of the level $-g/(1+\gamma)$ by the velocity after time τ_a^V and hence, $\tau_y^V > \zeta$. Using this observation and the strong Markov property at τ_b^V, we obtain

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\sup_{[0,\zeta]} V_t \geq y\right) = \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_b^V < \tau_a^V\right)\, \mathbb{P}_{(0,b)}\left(\tau_y^V < \tau_{-\frac{g}{1+\gamma}}^V\right). \tag{2.4.14}$$

The lower bound in the theorem follows from (2.4.12) and (2.4.14), with $C = -\log \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_b^V < \tau_a^V\right)$. $C > 0$ is finite via arguments similar to those of (2.4.13). $\square$

2.4.2 Gap bounds

To bound tail probabilities of H, we split the path of the process by hitting times of the velocity. Doing so allows the gap to be controlled by Brownian motion with drift and its running maximum.

Lemma 2.4.4. *For any $x > 0$,*

$$\sup_{\nu\in\left(-g/\gamma,\frac{\gamma x}{4}-g/\gamma\right]} \mathbb{P}_{(x/2,\nu)}\left(\tau_x^H < \sigma(0)\right) \leq \exp\{-xg/2\gamma\}$$

and

$$\inf_{h\geq(\gamma/2g)\log(2),\, \nu>-g/\gamma} \mathbb{P}_{(h,\nu)}\left(\tau_x^H < \sigma(0)\right) \geq \exp\{-2x\, g/\gamma\}.$$

Proof. Take any $x > 0$. On $t < \tau_x^H \wedge \sigma(0)$, with $(H_0, V_0) = (h, \nu)$, H_t is dominated by $h + \frac{\nu}{\gamma} + \frac{g}{\gamma^2} - B_t - tg/\gamma$, using (2.1.4). For the upper bound, use $h = x/2$. The function $s(x) = e^{2\frac{g}{\gamma}x}$ makes $s(z - B_t - \frac{g}{\gamma}t)$ a martingale for any z. By the optional stopping theorem,

$$\begin{aligned}\mathbb{P}_{(x/2,\nu)}\left(\tau_x^H < \sigma(0)\right) &\leq \mathbb{P}\left(x/2 + \frac{\nu}{\gamma} + \frac{g}{\gamma^2} - B_t - \frac{g}{\gamma}t \text{ hits } x \text{ before } 0\right) \\ &= \frac{s\left(x/2 + \frac{\nu}{\gamma} + \frac{g}{\gamma^2}\right) - 1}{s\,(x) - 1} \leq \exp\{-xg/2\gamma\},\end{aligned}$$

for all $\nu \leq \frac{\gamma x}{4} - \frac{g}{\gamma}$. To prove the lower bound first note that if $h > x$, then $\tau_x^H < \sigma(0)$. For $h \in [(\gamma/2g)\log(2),\, x]$ and $t \leq \sigma(0)$, noting $\nu - V_t \geq 0$ (by (2.1.5)), we use (2.1.4) to calculate,

$$H_t = h + \frac{\nu}{\gamma} - \frac{V_t}{\gamma} - B_t - \frac{g}{\gamma}t \geq (\gamma/2g)\log 2 - B_t - \frac{g}{\gamma}t.$$

The optional stopping theorem again gives,

$$\begin{aligned}\mathbb{P}_{(h,\nu)}\left(\tau_x^H < \sigma(0)\right) \geq \mathbb{P}\left((\gamma/2g)\log 2 - B_t - \frac{g}{\gamma}t \text{ hits } x \text{ before } 0\right) &= \frac{s\left((\gamma/2g)\log 2\right) - 1}{s\,(x) - 1} \\ &= \frac{\exp\{-2x\, g/\gamma\}}{1 - \exp\{-2x\, g/\gamma\}} \geq \exp\{-2x\, g/\gamma\}.\end{aligned}$$

□

Lemma 2.4.5. *Fix $a > -g/\gamma$. For any $x > 0$,*

$$\sup_{\nu\in\left[a,\left(a+\frac{g}{\gamma}\right)e^{\gamma^2 x/(4g)}-\frac{g}{\gamma}\right]} \mathbb{P}_{(x/2,\nu)}\left(\sigma(0) < \tau_a^V\right) \leq \frac{2\sqrt{2\gamma}}{\sqrt{\pi g x}} \exp\{-x\, g/8\gamma\}.$$

Proof. Observe that for any $x > 0$ and any $\nu > -g/\gamma$, when $(H_0, V_0) = (x/2, \nu)$ we obtain from (2.1.1),

$$H_t = H_0 + S_t - B_t + L_t \geq \frac{x}{2} - \frac{g}{\gamma}t - B_t,$$

where we used $S_t = \int_0^t V_u du \geq -gt/\gamma$ for all $t \geq 0$. From this bound, we conclude that $H_t \geq \frac{x}{4} - B_t$ for all $t \leq \frac{\gamma x}{4g}$. Thus, if $B_t < x/4$ for all $t \leq \frac{\gamma x}{4g}$, then $\sigma(0) > \frac{\gamma x}{4g}$. In particular, along with (2.1.4), this implies that if $t \leq \frac{\gamma x}{4g}$, $V_t = (\nu + g/\gamma)e^{-\gamma t} - g/\gamma$. The right hand side of this equation equals a when $t = \gamma^{-1}\log\left(\frac{\nu+\frac{g}{\gamma}}{a+\frac{g}{\gamma}}\right)$. If $\nu \in \left[a, \left(a + \frac{g}{\gamma}\right)e^{\gamma^2 x/(4g)} - \frac{g}{\gamma}\right]$, then $\gamma^{-1}\log\left(\frac{\nu+\frac{g}{\gamma}}{a+\frac{g}{\gamma}}\right) \leq \frac{\gamma x}{4g}$. We conclude that for $x > 0$ and $\nu \in \left[a, \left(a + \frac{g}{\gamma}\right)e^{\gamma^2 x/(4g)} - \frac{g}{\gamma}\right]$, if $B_t < x/4$ for all $t \leq \frac{\gamma x}{4g}$, then $\tau_a^V < \sigma(0)$. Consequently,

$$\sup_{\nu\in\left[a,\left(a+\frac{g}{\gamma}\right)e^{\gamma^2 x/(4g)}-\frac{g}{\gamma}\right]} \mathbb{P}_{(x/2,\nu)}\left(\sigma(0) < \tau_a^V\right) \leq \mathbb{P}\left(\sup_{t\leq\frac{\gamma x}{4g}} B_t \geq x/4\right) \leq \frac{2\sqrt{2\gamma}}{\sqrt{\pi g x}}\exp\left\{-xg/8\gamma\right\}.$$

□

Theorem 2.4.6. *There exist positive constants $x'(\gamma, g), C, C'$ such that*

$$e^{-C'}\exp\left\{-2xg/\gamma\right\} \leq \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H < \zeta\right) \leq e^{C}\ \exp\left\{-xg/16\gamma\right\},$$

for all $x > x'(\gamma, g)$.

Proof. Theorem 2.4.3 shows there exist positive constants $x_0(\gamma, g)$, C, C' such that

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau^V_{\sqrt{x\frac{g}{\gamma(1+\gamma)}}} < \zeta\right) \leq e^{C}\, e^{-\frac{1+\gamma}{4}\left(\sqrt{x\frac{g}{\gamma(1+\gamma)}}+g/(1+\gamma)\right)^2} \leq e^{C'}\, e^{-xg/4\gamma}, \tag{2.4.15}$$

for $x > x_0(\gamma, g)$. The union bound then gives,

$$\begin{aligned}\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H < \zeta\right) &\leq \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H < \zeta < \tau^V_{\sqrt{x\frac{g}{\gamma(1+\gamma)}}}\right) + \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau^V_{\sqrt{x\frac{g}{\gamma(1+\gamma)}}} < \zeta\right)\\ &\leq \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H < \zeta < \tau^V_{\sqrt{x\frac{g}{\gamma(1+\gamma)}}}\right) + e^{C}\, e^{-xg/4\gamma}.\end{aligned} \tag{2.4.16}$$

Fix $a = -\left(g + g/2\gamma\right)/(1+\gamma)$ and $b = -\left(g - g/2(1+\gamma)\right)/(1+\gamma)$, as in the definition (2.1.7) of ζ. Choose $x_0(\gamma, g)$ large enough that $\left(a + \frac{g}{\gamma}\right) e^{\gamma^2 x/(4g)} - \frac{g}{\gamma} > \frac{\gamma x}{4} - \frac{g}{\gamma} > \sqrt{x \frac{g}{\gamma(1+\gamma)}}$ for all $x > x_0(\gamma, g)$. The strong Markov property at $\tau^H_{x/2}$ and Lemmas 2.4.4 and 2.4.5 show there exist constants $x_1(\gamma, g) > x_0(\gamma, g)$ and $C > 0$ such that for $x > x_1(\gamma, g)$,

$$\sup_{\nu \in [a,0]} \mathbb{P}_{(0,\nu)}\left(\tau^H_x \leq \tau^V_a \wedge \tau^V_{\sqrt{x\frac{g}{\gamma(1+\gamma)}}}\right) \leq \sup_{\nu \in \left[a, \sqrt{x\frac{g}{\gamma(1+\gamma)}}\right]} \mathbb{P}_{(x/2,\nu)}\left(\tau^H_x \leq \tau^V_a\right)$$
$$\leq \sup_{\nu \in \left[a, \left(a+\frac{g}{\gamma}\right)e^{\gamma^2 x/(4g)} - \frac{g}{\gamma}\right]} \mathbb{P}_{(x/2,\nu)}\left(\sigma(0) < \tau^H_x \leq \tau^V_a\right) + \sup_{\nu \in \left[a, \frac{\gamma x}{4} - \frac{g}{\gamma}\right]} \mathbb{P}_{(x/2,\nu)}\left(\tau^H_x < \sigma(0)\right)$$
$$\leq e^C e^{-xg/8\gamma}. \quad (2.4.17)$$

Fix $x > x_1(\gamma, g)$. We define slight modifications of the stopping times given in (2.3.1), and therefore we use the same notation. Define $\alpha_{-1} = 0$ and $\alpha_0 = \tau^V_a \wedge \tau^V_b$. If $\tau^V_b < \tau^V_a$, define $\alpha_j = \alpha_0$ for all $k \geq 0$ and $N^- = 0$. For $k \geq 0$, if $V_{\alpha_{3k}} = a$,

$$\alpha_{3k+1} = \inf\left\{t \geq \alpha_{3k} \mid H_t \leq x/2\right\},$$
$$\alpha_{3k+2}, \alpha_{3k+3} \quad \text{defined exactly as in (2.3.1).} \quad (2.4.18)$$

If $V_{\alpha_{3k}} = -g/(1+\gamma)$ then $\alpha_j = \alpha_{3k}$ for all $j \geq 3k$. As before, set $N^- = \inf\left\{k \geq 1 \mid V_{\alpha_{3k}} = -g/(1+\gamma)\right\}$. We consider an arbitrary, fixed $x > x_1(\gamma, g)$ and suppress the dependence of $\alpha_1, \alpha_2 \ldots$ on x. Since $b < 0 < \sqrt{x\frac{g}{\gamma(1+\gamma)}}$, (2.4.17) shows,

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau^H_x \in (\alpha_{-1}, \alpha_0]\right) = \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau^H_x \leq \tau^V_a \wedge \tau^V_b\right)$$
$$\leq \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau^H_x \leq \tau^V_a \wedge \tau^V_{\sqrt{x\frac{g}{\gamma(1+\gamma)}}}\right) \leq e^C e^{-xg/8\gamma}. \quad (2.4.19)$$

Fix $k \geq 1$. Recall that on $N^- \geq k$, $(H_{\alpha_{3k-1}}, V_{\alpha_{3k-1}}) = (0, -(g + g/4\gamma)/(1+\gamma))$. Starting from α_{3k-1}, the velocity cannot rise to $\sqrt{x\frac{g}{\gamma(1+\gamma)}}$, for any $x > 0$, without first passing through $(0, -g/(1+\gamma))$, by (2.1.5). In addition, When $N < k$, $\alpha_{3k-1} = \alpha_{3k}$ and $\tau^H_x \in (\alpha_{3k-1}, \alpha_{3k}]$ is impossible. We use these observations, along with the strong Markov property at α_{3k-1} and

(2.4.17), to obtain for any $k \geq 1$,

$$\begin{aligned}
\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H \in (\alpha_{3k-1}, \alpha_{3k}]\right) &= \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H \in (\alpha_{3k-1}, \alpha_{3k}],\ N^- \geq k\right) \\
&= \mathbb{E}_{(0,-g/(1+\gamma))}\left(\mathbb{1}_{N^- \geq k}\, \mathbb{P}_{(H_{\alpha_{3k-1}}, V_{\alpha_{3k-1}})}\left(\tau_x^H \leq \tau_a^V \wedge \tau_{-g/(1+\gamma)}^V\right)\right) \\
&\leq \sup_{\nu \in [a,0]} \mathbb{P}_{(0,\nu)}\left(\tau_x^H \leq \tau_a^V \wedge \tau_{-g/(1+\gamma)}^V\right) \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^- \geq k\right) \\
&\leq e^C\, e^{-xg/8\gamma}\, \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^- \geq k\right). \quad (2.4.20)
\end{aligned}$$

Now we estimate the probability of $\tau_x^H \in (\alpha_{3k}, \alpha_{3k+1}]$ for $k \geq 0$. If $H_{\alpha_{3k}} \leq x/2$ then $\alpha_{3k+1} = \alpha_{3k}$ and $\tau_x^H \in (\alpha_{3k}, \alpha_{3k+1}]$ is impossible. Therefore, we need only consider cases in which H reaches $x/2$ between α_{3k-1}, when H is zero, and α_{3k}. Therefore, once again we apply the strong Markov property at α_{3k-1} and (2.4.17), with $x/2$ in place of x, to obtain

$$\begin{aligned}
&\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H \in (\alpha_{3k}, \alpha_{3k+1}]\right) \\
&\quad = \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H \in (\alpha_{3k}, \alpha_{3k+1}],\ \inf\{t \geq \alpha_{3k-1} \mid H_t \geq x/2\} \in (\alpha_{3k-1}, \alpha_{3k}],\ N^- \geq k\right) \\
&\quad\quad \leq \sup_{\nu \in [a,0]} \mathbb{P}_{(0,\nu)}\left(\tau_{x/2}^H \leq \tau_a^V \wedge \tau_{-g/(1+\gamma)}^V\right) \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^- \geq k\right) \\
&\quad\quad \leq e^C\, e^{-xg/16\gamma}\, \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^- \geq k\right). \quad (2.4.21)
\end{aligned}$$

Note that for any $k \geq 0$, the fact (2.1.5) that the velocity increases only where $H = 0$ and the definition of α_{3k+1} show $(H_{\alpha_{3k+1}}, V_{\alpha_{3k+1}}) \in [0, x/2] \times (-g/\gamma, a]$. Suppose we have initial conditions such that $(H_0, V_0) \in [0, x/2] \times (-g/\gamma, a]$. When $t < \tau^V_{-(g+g/4\gamma)/(1+\gamma)}$, noting $a < -(g+g/4\gamma)/(1+\gamma)$, we use (2.1.1), $S_t \leq -t(g+g/4\gamma)/(1+\gamma)$ and $-V_t > 0$ to show

$$H_t \leq H_t - V_t = H_0 - V_0 + (1+\gamma)S_t + tg - B_t \leq x/2 + g/\gamma - tg/4\gamma - B_t. \quad (2.4.22)$$

By (2.4.22), if H hits x before V hits $-(g+g/4\gamma)/(1+\gamma)$, then $\sup_{t<\infty}(-B_t - tg/4\gamma)$ must have reached $x/2 - g/\gamma$, which is positive so long as we have chosen $x_1(\gamma, g)$ large enough. We use the strong Markov property at α_{3k+1}, (2.4.22) and $\sup\limits_{t<\infty}(-B_t - tg/4\gamma) \overset{d}{=} \text{Exponential}(g/2\gamma)$

(see Chapter 3.5 of Karatzas and Shreve (1991)) to show for any $k \geq 0$,

$$
\begin{aligned}
\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H \in (\alpha_{3k+1}, \alpha_{3k+2}]\right) &= \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H \in (\alpha_{3k+1}, \alpha_{3k+2}],\ N^- \geq k+1\right) \\
&\leq \sup_{(h,\nu)\in[0,x/2]\times(-g/\gamma,a]} \mathbb{P}_{(h,\nu)}\left(\tau_x^H < \tau^V_{-(g+g/4\gamma)/(1+\gamma)}\right) \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^- \geq k+1\right) \\
&\leq \mathbb{P}\left(\sup_{t<\infty}(-B_t - tg/4\gamma) \geq x/2 - g/\gamma\right) \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^- \geq k\right) \\
&= e^{-(g/2\gamma)(x/2-g/\gamma)}\, \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^- \geq k\right) = e^C\, e^{-xg/4\gamma}\, \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^- \geq k\right). \quad (2.4.23)
\end{aligned}
$$

Recall that when $\tau_a^V < \tau_b^V$, we have $\zeta = \alpha_{3N^-}$. Therefore, on $\tau_a^V < \tau_b^V$ we have $\tau_x^H < \zeta$ if and only if $\tau_x^H \in (\alpha_{3k+j}, \alpha_{3k+j+1}]$ for some $k \geq 0$, $j = -1, 0, 1$. As a result, we use (2.4.19), (2.4.20), (2.4.21) and (2.4.23) to show

$$
\begin{aligned}
\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H < \zeta,\ \tau_a^V < \tau_b^V\right) &= \sum_{k=0}^{\infty}\sum_{j=-1}^{1} \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H \in (\alpha_{3k+j}, \alpha_{3k+j+1}]\right) \\
&\leq 3e^C\, e^{-xg/16\gamma} \sum_{k=0}^{\infty} \mathbb{P}_{(0,-g/(1+\gamma))}\left(N^- \geq k\right). \quad (2.4.24)
\end{aligned}
$$

The exact same argument as in Lemma 2.3.1 shows $\mathbb{E}_{(0,-g/(1+\gamma))}\, N^- \leq e^{C'}$. Recalling that $\tau_a^V < \tau_b^V$ implies $\zeta < \tau^V_{\sqrt{x\frac{g}{\gamma(1+\gamma)}}}$, we have by (2.4.24) a positive constant C'' such that,

$$
\begin{aligned}
\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H < \zeta < \tau^V_{\sqrt{x\frac{g}{\gamma(1+\gamma)}}},\ \tau_a^V < \tau_b^V\right) &= \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H < \zeta,\ \tau_a^V < \tau_b^V\right) \\
&\leq e^{C''}\, e^{-xg/16\gamma}. \quad (2.4.25)
\end{aligned}
$$

Now consider the case $\tau_b^V < \tau_a^V$. Once again fix $x > x_1(\gamma, g)$. (2.4.17) directly implies

$$
\begin{aligned}
\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H \leq \tau_b^V < \tau_a^V\right) &\leq \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H \leq \tau_a^V \wedge \tau^V_{\sqrt{x\frac{g}{\gamma(1+\gamma)}}}\right) \\
&\leq e^C\, e^{-xg/8\gamma}. \quad (2.4.26)
\end{aligned}
$$

We now control H in the time between τ_b^V and ζ. There are two possibilities: Either $\zeta < \tau_a^V$, or V crosses down to a before the renewal time is reached. In the former case: Since $b \in [a, 0]$

and $H_{\tau_b^V} = 0$ by (2.1.5), we use the strong Markov property at τ_b^V and (2.4.17) again to show

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H \in (\tau_b^V, \zeta],\ \zeta < \tau_a^V \wedge \tau^V_{\sqrt{x\frac{g}{\gamma(1+\gamma)}}}\right) \le \mathbb{P}_{(0,b)}\left(\tau_x^H \le \tau_a^V \wedge \tau^V_{\sqrt{x\frac{g}{\gamma(1+\gamma)}}}\right)$$
$$\le e^C\, e^{-xg/8\gamma}. \quad (2.4.27)$$

In the case where $\tau_a^V < \zeta$, we modify the analysis used to prove (2.4.25), as follows. Define $\tilde{\beta}_{-1} = \tau_a^V \wedge \tau_b^V$ and $\tilde{\beta}_0 = \inf\{t \ge \tilde{\beta}_{-1} \,|\, V_t = a\}$. Define $\{\tilde{\beta}_j\}_{j\ge1}$ and $\tilde{N}^-$ analogously to $\{\alpha_j\}_{j\ge1}$ and N^- in (2.4.18). The strong Markov property at $\tilde{\beta}_{-1}$ and (2.4.17) show,

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H \in (\tilde{\beta}_{-1}, \tilde{\beta}_0],\ \tau_a^V < \zeta < \tau^V_{\sqrt{x\frac{g}{\gamma(1+\gamma)}}}\right) \le \mathbb{P}_{(0,b)}\left(\tau_x^H \le \tau_a^V \wedge \tau^V_{\sqrt{x\frac{g}{\gamma(1+\gamma)}}}\right)$$
$$\le e^C\, e^{-xg/8\gamma}. \quad (2.4.28)$$

The analysis of (2.4.19), (2.4.20), (2.4.21) and (2.4.23) is now repeated, with $\tilde{\beta}_j$ in place of α_j for $j \ge 0$ and $\tilde{N}^-$ in place of N^-, giving for $k \ge 0$ and $j = 0, 1, 2$,

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H \in (\tilde{\beta}_{3k+j}, \tilde{\beta}_{3k+j+1}],\ \tau_a^V < \zeta < \tau^V_{\sqrt{x\frac{g}{\gamma(1+\gamma)}}}\right)$$
$$\le e^C\, e^{-xg/16\gamma}\, \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tilde{N}^- \ge k\right). \quad (2.4.29)$$

Combining (2.4.26), (2.4.27), (2.4.28) and (2.4.29),

$$\begin{aligned}
&\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H < \zeta < \tau^V_{\sqrt{x\frac{g}{\gamma(1+\gamma)}}},\ \tau_b^V < \tau_a^V\right) \\
&\qquad \le \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H \in (0,\tau_b^V],\ \zeta < \tau^V_{\sqrt{x\frac{g}{\gamma(1+\gamma)}}},\ \tau_b^V < \tau_a^V\right) \\
&\qquad\quad + \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H \in (\tau_b^V,\zeta],\ \zeta < \tau_a^V \wedge \tau^V_{\sqrt{x\frac{g}{\gamma(1+\gamma)}}}\right) \\
&\qquad\quad + \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H \in (\tau_b^V,\zeta],\ \tau_a^V < \zeta < \tau^V_{\sqrt{x\frac{g}{\gamma(1+\gamma)}}}\right) \\
&\qquad \le 2e^{C}\, e^{-xg/8\gamma} + \sum_{k=0}^{\infty}\sum_{j=-1}^{1} \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H \in (\tilde\beta_{3k+j},\tilde\beta_{3k+j+1}]\right) \\
&\qquad \le 2e^{C}\, e^{-xg/8\gamma} + e^{C}\, e^{-xg/16\gamma}\sum_{k=0}^{\infty}\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tilde N^- \ge k\right). \qquad (2.4.30)
\end{aligned}$$

Arguing as in (2.4.25), we achieve,

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H < \zeta < \tau^V_{\sqrt{x\frac{g}{\gamma(1+\gamma)}}},\ \tau_b^V < \tau_a^V\right) \le e^{C''}\, e^{-xg/16\gamma}. \qquad (2.4.31)$$

Our choice of $x > x_1(\gamma, g)$ in (2.4.16), (2.4.25) and (2.4.31) was arbitrary, so the upper bound of the theorem is proven.

We now prove the lower bound. For any $x > 0$, we consider a path in which H attains a positive value h_0 before the velocity leaves $[a, b]$, then rises to x before returning zero. Since by definition (2.1.7), $\zeta > \tau_a^V \wedge \tau_b^V$ and $H_\zeta = 0$, this implies $\tau_x^H < \zeta$. We select any $h_0 \ge (\gamma/2g)\log 2$ and $x > h_0$. The strong Markov property at $\tau_{h_0}^H$ and Lemma 2.4.4 give,

$$\begin{aligned}
\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_x^H < \zeta\right) &\ge \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_{h_0}^H < \tau_a^V \wedge \tau_b^V,\ \tau_{h_0}^H < \tau_x^H < \sigma(\tau_{h_0}^H)\right) \\
&\ge \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_{h_0}^H < \tau_a^V \wedge \tau_b^V\right) \inf_{\nu\in(a,b)} \mathbb{P}_{(h_0,\nu)}\left(\tau_x^H < \sigma(0)\right) \ge e^{-C}\, e^{-2xg/\gamma}, \qquad (2.4.32)
\end{aligned}$$

where $C = -\log\left(\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_{h_0}^H < \tau_a^V \wedge \tau_b^V\right)\right)$. It remains only to show $C < \infty$. Suppose $H_0 = 0$ and $V_0 = -g/(1+\gamma)$. Define $T = \frac{1}{\gamma}\log\left(\frac{-g/(1+\gamma)+g/\gamma}{a+g/\gamma}\right) = \frac{1}{\gamma}\log 2$. By (2.1.6), V cannot hit a before time T and hence, $S_u \ge ua$ for $u < T$. Thus, if $\inf_{u<T} B_u < -(h_0 - Ta) < 0$

(recalling $a < 0$), we obtain from system equations (2.1.1)

$$\sup_{u<T} H_u = \sup_{u<T}(S_u - B_u + L_u) \geq \sup_{u<T} (ua - B_u) > h_0.$$

If, in addition, $\sup_{u<T} (B_u - ua) < b + g/(1+\gamma)$ then $V_u \leq -g/(1+\gamma) - u(\gamma a + g) + \sup_{s<T} (B_s - sa) < b$ for all $u < T$, since $\gamma a + g > 0$. Recalling that $b + g/(1+\gamma) > 0$, we have shown

$$\begin{aligned} &\mathbb{P}_{(0,-g/(1+\gamma))} \left(\tau^H_{h_0} < \tau^V_a \wedge \tau^V_b\right) \\ &\qquad \geq \mathbb{P}\left(\inf_{u<T} B_u < -(h_0 - Ta),\ \sup_{u<T} (B_u - ua) < b + g/(1+\gamma)\right) > 0. \end{aligned} \tag{2.4.33}$$

(2.4.33) shows $C < \infty$ in (2.4.32), and the theorem is proven with $x'(\gamma, g) = x_1(\gamma, g) \vee [(\gamma/2g)\log 2]$. $\square$

2.5 Tail bounds for π and path fluctuations

This section is devoted to the proof of Theorems 2.1.3 and 2.1.4.

Proof of Theorem 2.1.3. First we prove the theorem's lower bound for $\pi(\mathbb{R}_+ \times (y, \infty))$. Fix $y'(\gamma, g)$ as in Theorem 2.4.3 and $y > y'(\gamma, g) > 0$. Recall the notation $\tau^V_z(\alpha) = \inf\{t \geq \alpha : V_t = z\}$ for any stopping time α. On the set $\tau^V_{2y} < \zeta$, we have $\tau^V_y(\tau^V_{2y}) < \zeta$. Theorem 2.4.3, the strong Markov property at τ^V_y and the definition of π in Theorem 2.1.2 show there exists a constant $C > 0$ such that,

$$\begin{aligned} &\mathbb{E}_{(0,-g/(1+\gamma))}(\zeta)\ \pi(\mathbb{R}_+ \times (y,\infty)) = \mathbb{E}_{(0,-g/(1+\gamma))}\left(\int_0^{\zeta} \mathbb{1}_{V_t > y}\, dt\right) \\ &\geq \mathbb{E}_{(0,-g/(1+\gamma))}\left(\mathbb{1}_{\tau^V_{2y}<\zeta} \int_{\tau^V_{2y}}^{\tau^V_y(\tau^V_{2y})} \mathbb{1}_{V_t>y}\, dt\right) = \mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau^V_{2y} < \zeta\right)\ \mathbb{E}_{(0,2y)}\left(\int_0^{\tau^V_y} \mathbb{1}_{V_t>y}\, dt\right) \\ &\qquad \geq e^{-C}\, e^{-2(1+\gamma)(2y+g/(1+\gamma))^2}\, \mathbb{E}_{(0,2y)}\, \tau^V_y \geq e^{-C}\, e^{-4(1+\gamma)(y+g/(1+\gamma))^2}\, \mathbb{E}_{(0,2y)}\, \tau^V_y. \end{aligned} \tag{2.5.1}$$

By (2.1.6), $\mathbb{E}_{(0,2y)}\, \tau^V_y \geq \frac{1}{\gamma}\log\left(\frac{2y+g/\gamma}{y+g/\gamma}\right) \geq \frac{1}{\gamma}\log\left(\frac{2y'(\gamma,g)+g/\gamma}{y'(\gamma,g)+g/\gamma}\right) > 0$ for all $y > y'(\gamma, g)$. So (2.5.1) proves the lower bound of the theorem for $\pi(\mathbb{R}_+ \times (y,\infty))$, for all $y > y'(\gamma, g)$.

We now prove the lower bound for $\pi\left((x,\infty)\times(-g/\gamma,\infty)\right)$. Fix $x'(\gamma,g)>0$ as in Theorem 2.4.6 and $x>x'(\gamma,g)$. Proceeding similarly to (2.5.1), by Theorem 2.4.6 there exists a $C'>0$ such that,

$$\mathbb{E}_{(0,-g/(1+\gamma))}\,(\zeta)\;\pi\left((x,\infty)\times(-g/\gamma,\infty)\right)\geq e^{-C'}\,e^{-4xg/\gamma}\inf_{\nu>-g/\gamma}\mathbb{E}_{(2x,\nu)}\,\tau_x^H. \tag{2.5.2}$$

When $H_0=2x$, (2.1.1) and (2.1.3) show $H_t\geq 2x+S_t-B_t\geq 2x-tg/\gamma-B_t$ for any $V_0>-g/\gamma$. Therefore, $\tau_x^H\geq\inf\{t\geq 0\,:\,-B_t-tg/\gamma=-x\}$ for any initial condition V_0. The expected hitting time of Brownian motion with drift $-g/\gamma$ at level $-x'(\gamma,g)$ is strictly positive and finite (e.g. Ch. 3C in Karatzas and Shreve (1991)), so we have $C''>0$ such that $\inf_{\nu>-g/\gamma}\mathbb{E}_{(2x,\nu)}\,\tau_x^H\geq C''$ for each $x>x'(\gamma,g)$. This fact and (2.5.2) prove the required lower bound for $\pi\left((x,\infty)\times(-g/\gamma,\infty)\right)$.

We now show the upper bounds of the theorem. Again using the representation for π in Theorem 2.1.2 and the velocity bounds in Theorem 2.4.3, we obtain for $y>y'(\gamma,g)$,

$$\begin{aligned}\mathbb{E}_{(0,-g/(1+\gamma))}\,(\zeta)\;\pi\left(\mathbb{R}_+\times(y,\infty)\right)&=\mathbb{E}_{(0,-g/(1+\gamma))}\left(\int_0^\zeta \mathbb{1}_{V_t>y}\,dt\right)\\ \leq\mathbb{E}_{(0,-g/(1+\gamma))}\left(\mathbb{1}_{\tau_y^V<\zeta}\left(\zeta-\tau_y^V\right)\right)&\leq\sqrt{\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_y^V<\zeta\right)}\sqrt{\mathbb{E}_{(0,-g/(1+\gamma))}(\zeta^2)}\\ &\leq e^{c/2}\,e^{-\frac{1+\gamma}{8}(y+g/(1+\gamma))^2}\sqrt{\mathbb{E}_{(0,-g/(1+\gamma))}(\zeta^2)}.\end{aligned} \tag{2.5.3}$$

The upper bound for $\pi\left(\mathbb{R}_+\times(y,\infty)\right)$ follows from (2.5.3) upon noting that $\mathbb{E}_{(0,-g/(1+\gamma))}(\zeta^2)<\infty$, which is a consequence of Theorem 2.1.1. The upper bound for $\pi\left((x,\infty)\times(-g/\gamma,\infty)\right)$ is proven similarly using Theorem 2.4.6. □

Proof of Theorem 2.1.4. The argument is identical to the one provided for Theorem 2.2 of Banerjee et al. (2019) and Theorem 2.2 of Banerjee and Mukherjee (2019a), so we give it cursory treatment. To demonstrate the fluctuation result for V: Theorem 2.4.3, (2.1.8) and the Borel-Cantelli lemmas give for any $\epsilon\in(0,1)$,

$$\frac{\sqrt{1-\epsilon}}{\sqrt{2}\sqrt{1+\gamma}}\leq\limsup_{n\to\infty}\frac{\sup_{t\in[\zeta_n,\zeta_{n+1}]}V_t}{\sqrt{\log n}}\leq 2\frac{\sqrt{1+\epsilon}}{\sqrt{1+\gamma}},\qquad\text{almost surely.}$$

A sub-sequence argument, and the observation $\lim_{n\to\infty} \frac{\zeta_n}{n} = \mathbb{E}\zeta_0$ almost surely, complete the proof. The second statement is proven similarly, with Theorem 2.4.6. $\square$

Proof of Theorem 2.1.5. By Theorem 2.1.4,

$$0 = \lim_{t\to\infty} \frac{H_t - V_t}{t} = \lim_{t\to\infty} \frac{H_0 - V_0 + (1+\gamma)\, S_t - B_t + gt}{t}, \tag{2.5.4}$$

and thus, using $\frac{B_t}{t} \to 0$ almost surely as $t \to \infty$, we obtain $\frac{S_t}{t} \to -\frac{g}{1+\gamma}$. In addition, again by Theorem 2.1.4, $\lim_{t\to\infty} \frac{H_t}{t} = \lim_{t\to\infty} \frac{S_t - X_t}{t} = 0$, giving the result. $\square$

2.6 Exponential ergodicity

This section will prove Theorem 2.1.6. We first show the process is ergodic, in the sense that $P^t((h,\nu,\,\cdot\,)$ converges to π in total variation, using coupling techniques of Ch. 10 Thorisson (2000). We show this can be upgraded to exponentially fast convergence using Lyapunov function techniques.

Typical proofs of exponential ergodicity via Harris' Theorem available in the literature (e.g. Hairer and Mattingly (2011); Meyn and Tweedie (2009); Mattingly et al. (2012); Cooke et al. (2017); Budhiraja and Lee (2007)) rely on the existence of continuous densities of P^t with respect to Lebesgue measure. Such densities give a positive chance of coupling two versions of the process within a set toward which the process has a strong drift.

In our case the transition laws do not have densities, as can be verified by observing the velocity decreases deterministically away from the boundary ∂S. In addition, the generator of the process is *not hypoelliptic* in the interior of the domain, which makes the situation more complicated: Hypoellipticity is a standard tool for establishing exponential ergodicity in the absence of ellipticity (Mattingly and Stuart (2002)). We adapt techniques from Thorisson (2000), which involve coupling the renewal times of two versions of the process, to furnish exponential ergodicity for our model.

Before we proceed to details, we provide a proof outline. The first step is proving convergence to stationarity in total variation distance. This is done using Theorem 3.3, Ch. 10 of Thorisson (2000) after proving in Lemma 2.6.1 that the gaps between the renewal times are 'spread-out' (see Section 3.5, Ch. 10 of Thorisson (2000)). The total variation convergence result is stated in Theorem 2.6.2.

This convergence is then upgraded to exponential ergodicity by establishing the *drift condition* (2.6.24) and *minorization condition* (2.6.23) which can be used to furnish exponential ergodicity using the recipe in Down et al. (1995). The drift condition is established via Lyapunov functions that can be obtained from exponential moments of hitting times of certain carefully chosen sets. The finiteness of exponential moments for one such set Λ is shown in Lemma 2.6.4.

The minorization condition is obtained by establishing certain structural properties for the Markov process along the lines of Meyn and Tweedie (1993a), which are stated in Lemma 2.6.3. In particular, we show that the process $(H_t, V_t)_{t\geq 0}$ is π-irreducible and positive Harris recurrent and the aforementioned set Λ is 'petite'. These structural properties imply a stronger version of (2.6.23) as stated in Lemma 2.6.5. Finally, the proof of Theorem 2.1.6 is completed at the end of the Section.

Lemma 2.6.1. *There exists a non-negative function f with $\int_0^\infty f(x)\,dx > 0$ such that for every measurable set $A \subset [0,\infty)$,*

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\zeta \in A\right) \geq \int_A f(x)\,dx.$$

Proof. Recall a, b in the definition (2.1.7) of ζ, where $-g/\gamma < a < -g/(1+\gamma) < b < 0$. By definition of ζ and (2.1.5), when $\tau_a^V < \tau_b^V$ we have $\zeta = \tau_a^V + \inf\{t \geq 0 \,:\, V_{\tau_a^V + t} = -g/(1+\gamma)\}$. For any measurable $A \subset [0,\infty)$, we apply the strong Markov property at τ_a^V to obtain,

$$\mathbb{P}_{(0,-g/(1+\gamma))}\left(\zeta \in A\right) \geq \mathbb{P}_{(0,-g/(1+\gamma))}\left(\zeta \in A,\, \tau_a^V < \tau_b^V\right) = \mathbb{E}_{(0,-g/(1+\gamma))}\left(\mathbb{1}_{\tau_a^V < \tau_b^V}\, F_A\left(H_{\tau_a^V}, \tau_a^V\right)\right), \tag{2.6.1}$$

where $F_A(h,t) = \mathbb{P}_{(h,a)}\left(\tau^V_{-g/(1+\gamma)} \in A - t\right)$. We will show the right-hand side of (2.6.1) has a density with respect to Lebesgue measure λ on $[0,\infty)$. By the Radon-Nikodym Theorem, it suffices to show that for each $u > 0$ and $A \subset [0,u]$,

$$\lambda(A) = 0 \implies \mathbb{E}_{(0,-g/(1+\gamma))}\left(\mathbb{1}_{\tau_a^V < \tau_b^V}\, F_A\left(H_{\tau_a^V}, \tau_a^V\right)\right) = 0. \tag{2.6.2}$$

Fix such a u and A, $h \geq 0$, and set $(H_0, V_0) = (h, a)$. (2.1.5) says that if t is a point of increase of V_t it must be a point of increase for $L_t = L_t^{(h,a)} = 0 \vee \sup_{s \leq t} (-h + B_s - S_s)$. As a result,

$$\tau^V_{-g/(1+\gamma)} = \inf\{t \geq 0 \,:\, B_t - (1+\gamma)S_t - gt = h - g/(1+\gamma) - a\} = \tau^W_{h-g/(1+\gamma)-a}, \qquad (2.6.3)$$

where $W_t = B_t - \int_0^t (1+\gamma) V_{s \wedge \tau_0^V} - gt$. The second equality in (2.6.3) follows from the fact that $-g/(1+\gamma) < 0$ and $V_0 = a < -g/(1+\gamma)$, so $W_t = B_t - (1+\gamma)S_t - gt$ for $t \leq \tau^V_{-g/(1+\gamma)}$. Since $t \mapsto V_{t \wedge \tau_0^V}$ is bounded, the Novikov condition and Girsanov's theorem (Karatzas and Shreve (1991) Ch. 3.5) show the law of the process W is equivalent to that of standard Brownian motion on any bounded time interval $[0, u]$. Therefore, noting that $h - g/(1+\gamma) - a > 0$, $\tau^W_{h-g/(1+\gamma)-a}$ has a density with respect to Lebesgue measure (Karatzas and Shreve (1991) Ch. 2.8). Now (2.6.3) implies that whenever $\lambda(A) = 0$, $F_A(h, t) = 0$ for any $t \in [0, u]$ and $h \geq 0$. This proves (2.6.2). f can thus be taken as the density with respect to λ of the measure $A \mapsto \mathbb{E}_{(0,-g/(1+\gamma))}\left(\mathbb{1}_{\tau_a^V < \tau_b^V} F_A\left(H_{\tau_a^V}, \tau_a^V\right)\right)$. Since $\mathbb{P}_{(0,-g/(1+\gamma))}\left(\tau_a^V < \tau_b^V\right) > 0$, we have $\int_0^\infty f(x)\, dx > 0$. The lemma follows. $\square$

Theorem 2.6.2. *For every initial condition $(h, \nu) \in S$,*

$$\left\| P^t((h,\nu), \,\cdot\,) - \pi \right\|_{TV} \longrightarrow 0 \quad \text{as } t \to \infty.$$

Proof. This is a consequence of Lemma 2.6.1 and Theorem 3.3, Ch. 10 of Thorisson (2000). Using the terminology from that reference: The process (H, V) is classical regenerative by the strong Markov property and (2.1.7), (2.1.8). Since $\mathbb{P}_{(h,\nu)}(\zeta_1 - \zeta_0 \in \cdot) = \mathbb{P}_{(0,-g/(1+\gamma))}(\zeta \in \cdot)$, Lemma 2.6.1 shows the inter-regeneration time $\zeta_1 - \zeta_0$ is 'spread out.' The proof is completed by Theorem 3.3 (b), Ch. 10 of Thorisson (2000). $\square$

The following Lemma establishes several structural properties for the process $(H_t, V_t)_{t \geq 0}$. We refer the reader to Meyn and Tweedie (1993a) for definitions of resolvent kernels, π-irreducibility, 'petite' sets and Harris recurrence that appear in the following Lemma.

Lemma 2.6.3. *Define the (discrete time) resolvent transition kernel $R((h,\nu), \,\cdot\,) := \int_0^\infty e^{-t} P^t((h,\nu), \,\cdot\,)\, dt$, and the set $\Lambda = [0,1] \times \left[-\frac{g+g/2(1+\gamma)}{1+\gamma},\ g/\gamma\right]$. The following hold:*

(a) *For any measurable set A, $\pi(A) > 0$ implies $R((h,\nu),A) > 0$ for any $(h,\nu) \in S$.*

(b) *For any measurable set A, $\pi(A) > 0$ implies $\int_0^\infty P^t\left((h,\nu),\,\cdot\,\right)dt > 0$ for any $(h,\nu) \in S$. In other words, the process is π-irreducible.*

(c) *There exists an $\alpha > 0$ and a non-trivial measure μ, equivalent to π, such that $R((h,\nu),A) \geq \alpha\,\mu(A)$ for all $(h,\nu) \in \Lambda$ and all measurable A. In particular, Λ is a 'petite' set.*

(d) *The process $(H_t, V_t)_{t\geq 0}$ is Harris recurrent, and since the invariant measure π is finite the process is positive Harris recurrent.*

Proof. Using the definition of successive renewal times $\zeta := \zeta_0$ and $\{\zeta_n\}_{n\geq 0}$ given in (2.1.7), we apply the strong Markov property at $\zeta_n, n \geq 0$ to show for any measurable set A, with $\mathcal{F}_{\zeta_n}$ denoting the stopped filtration for (H,V),

$$\begin{aligned}\mathbb{E}_{(h,\nu)}\left(\int_{\zeta_n}^{\zeta_{n+1}} e^{-t}\mathbb{1}_{(H_t,V_t)\in A}\,dt \mid \mathcal{F}_{\zeta_n}\right) &= \int_{\zeta_n}^{\infty} e^{-t}\,\mathbb{E}_{(0,-g/(1+\gamma))}\left(\mathbb{1}_{\zeta>t-\zeta_n,(H_{t-\zeta_n},V_{t-\zeta_n})\in A}\right)dt \\ &= e^{-\zeta_n}\,\mathbb{E}_{(0,-g/(1+\gamma))}\left(\int_0^{\zeta} e^{-t}\mathbb{1}_{(H_t,V_t)\in A}\,dt\right). \qquad (2.6.4)\end{aligned}$$

Define $\mu(A) := \mathbb{E}_{(0,-g/(1+\gamma))}\left(\int_0^\zeta e^{-t}\mathbb{1}_{(H_t,V_t)\in A}\,dt\right)$. Taking expectations in (2.6.4) and summing over n we have,

$$\begin{aligned}R\left((h,\nu),\,A\right) &= \mathbb{E}_{(h,\nu)}\left(\int_0^\zeta e^{-t}\mathbb{1}_{(H_t,V_t)\in A}\,dt\right) + \sum_{n=0}^{\infty}\mathbb{E}_{(h,\nu)}\int_{\zeta_n}^{\zeta_{n+1}} e^{-t}\mathbb{1}_{(H_t,V_t)\in A}\,dt \\ &= \mathbb{E}_{(h,\nu)}\left(\int_0^\zeta e^{-t}\mathbb{1}_{(H_t,V_t)\in A}\,dt\right) + \mu(A)\sum_{n=0}^{\infty}\mathbb{E}_{(h,\nu)}\,e^{-\zeta_n}. \qquad (2.6.5)\end{aligned}$$

The form of π given in Theorem 2.1.2 shows the measures π and μ are equivalent. Thus, (2.6.5) proves (a), and (b) is an immediate consequence of (a).

The arguments used to prove Theorem 2.1.1 are valid for any starting point (h,ν) other than $(0,-g/(1+\gamma))$: Since the process started from (h,ν) can reach the points b, a used in (2.3.7), (2.3.1) in finite time, the analysis to show bounds as in Theorem 2.1.1 goes through for any initial conditions. Therefore, $\mathbb{E}_{(h,\nu)}e^{-\zeta} > 0$ for all (h,ν) and is bounded below on bounded

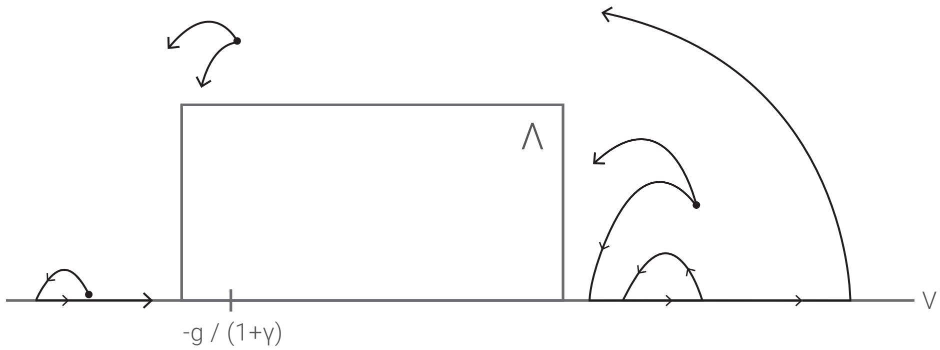

Figure 2.1: Process with initial conditions in left, top, right regions relative to Λ.

sets. Setting $\alpha = \inf_{(h,\nu)\in\Lambda} \sum_{n=0}^{\infty} \mathbb{E}_{(h,\nu)}\, e^{-\zeta_n}$, (2.6.5) gives (c). Note that setting $A = S$ in (2.6.5) shows $\alpha = \left(1 - \mathbb{E}_{(0,-g/(1+\gamma))}\, e^{-\zeta}\right)^{-1} \inf_{(h,\nu)\in\Lambda} \mathbb{E}_{(h,\nu)}\, e^{-\zeta}$.

Since, as noted in the previous paragraph, $\mathbb{P}_{(h,\nu)}(\zeta < \infty) = 1$ for all $(h,\nu) \in S$ and $(0, -g/(1+\gamma))$ is contained in Λ, we have $\mathbb{P}_{(h,\nu)}(\tau_\Lambda < \infty) = 1$. Moreover, part (c) of the lemma shows that the set Λ is *petite* in the sense of Section 4 in Meyn and Tweedie (1993a). These two observations imply (d) by Theorem 4.3 (ii) of Meyn and Tweedie (1993a). $\square$

Lemma 2.6.4. *For Λ as in Lemma 2.6.3, there exists a continuous function F such that,*

$$\mathbb{E}_{(h,\nu)}\, e^{\eta\tau_\Lambda(1)} \le F(h,\nu), \quad \text{for all } \eta < \frac{1}{16}(g/(1+\gamma))^2,\ (h,\nu) \in S, \tag{2.6.6}$$

where $\tau_\Lambda(1) := \{t \ge 1 : (H_t, V_t) \in \Lambda\}$. In particular, the function $\mathbb{E}_{(h,\nu)}\, e^{\eta\tau_\Lambda(1)}$ is finite for all (h,ν) and bounded on Λ, uniformly for $\eta < \frac{1}{16}(g/(1+\gamma))^2$.

Proof. Throughout the proof, the terms $c, c' > 0$ will denote constants dependent only on g, γ, not η, whose values might change from line to line. We will define a range of η to satisfy the theorem, giving upper bounds on such admissible η as needs arise.

First we bound $\mathbb{E}_{(h,\nu)}\, e^{\eta\tau_\Lambda}$ for $(h,\nu) \notin \Lambda$, recalling $\tau_\Lambda = \inf\{t \ge 0 : (H_t, V_t) \in \Lambda\}$. We consider three regions for initial conditions (h,ν): the left, the top and the right of the box Λ.

Define $\tilde{a} = -(g + g/2(1+\gamma))/(1+\gamma)$. Consider initial conditions to the left of Λ, that is $(H_0, V_0) = (h, \nu)$ for $h \geq 0$ and $\nu \in (-g/\gamma, \tilde{a})$. Recall (2.1.5) says V increases only on the set $\{s : H_s = 0\}$. This implies $\tau_{\tilde{a}}^V = \tau_\Lambda$. Lemma .0.6 with $\epsilon_0 = g/2(1+\gamma)$ and $u = \tilde{a}$ show,

$$\sup_{\nu \in (-g/\gamma, \tilde{a})} \mathbb{E}_{(h,\nu)}\, e^{\eta \tau_\Lambda} \leq e^c e^{4\eta(1+\gamma)h/g}, \quad h \geq 0,\ \nu \in (-g/\gamma, \tilde{a}),\ \eta < \frac{1}{16}\left(g/(1+\gamma)\right)^2. \qquad (2.6.7)$$

Consider $V_0 = \tilde{a}$ and $H_0 > 1$. (2.1.5) again shows $\tau_{\tilde{a}}^{V+} = \inf\{t > 0 : V_t = \tilde{a}\} = \tau_\Lambda$. Remark .0.1 after Lemma .0.6 gives,

$$\mathbb{E}_{(h,\tilde{a})}\, e^{\eta \tau_\Lambda} \leq e^c e^{4\eta(1+\gamma)h/g}, \quad h > 1,\ \eta < \frac{1}{16}\left(g/(1+\gamma)\right)^2. \qquad (2.6.8)$$

Now take $h > 1, \nu \in (\tilde{a}, \frac{g}{\gamma}]$, that is, initial conditions above Λ. Since V cannot increase before H hits zero, there are only two ways the process can enter Λ as depicted in Figure 2.1: Through the top of the rectangle, when $\tau_1^H \leq \tau_{\tilde{a}}^V$, or if the velocity drops below $\tilde{a}$ and the process enters Λ when V next rises to hit $\tilde{a}$. (2.1.4) shows $\left\{\tau_1^H \leq \tau_{\tilde{a}}^V,\ \tau_1^H > \frac{1}{\gamma}\log\left(\frac{\nu+g/\gamma}{\tilde{a}+g/\gamma}\right)\right\} = \varnothing$. As a result,

$$\mathbb{E}_{(h,\nu)}\left(e^{\eta\,\tau_\Lambda} \mathbb{1}_{\tau_1^H \leq \tau_{\tilde{a}}^V}\right) = \mathbb{E}_{(h,\nu)}\left(e^{\eta\,\tau_1^H} \mathbb{1}_{\tau_1^H \leq \tau_{\tilde{a}}^V}\right) \leq \left(\frac{\nu + g/\gamma}{\tilde{a} + g/\gamma}\right)^{\eta/\gamma} \leq e^c, \qquad (2.6.9)$$

for $\tilde{a} < \nu \leq g/\gamma$ and $\eta < \frac{1}{16}\left(g/(1+\gamma)\right)^2$. In addition, for the same range on ν, η, using the strong Markov property at $\tau_{\tilde{a}}^V$ and (2.6.8),

$$\mathbb{E}_{(h,\nu)}\left(e^{\eta\,\tau_\Lambda} \mathbb{1}_{\tau_{\tilde{a}}^V < \tau_1^H}\right) \leq e^c\, \mathbb{E}_{(h,\nu)}\left(\mathbb{1}_{\tau_{\tilde{a}}^V < \tau_1^H} e^{(4\eta(1+\gamma)/g) H_{\tau_{\tilde{a}}^V}}\right). \qquad (2.6.10)$$

Now (2.1.4) implies $H_t \leq h + c' + \sup_{t<\infty}\left(-B_t - t\frac{g}{\gamma}\right)$ for $t < \tau_{\tilde{a}}^V \wedge \tau_1^H$. Since $\sup_{t<\infty}\left(-B_t - t\frac{g}{\gamma}\right)$ is exponentially distributed with mean $\frac{\gamma}{2g}$, (2.6.10) gives,

$$\mathbb{E}_{(h,\nu)}\left(e^{\eta\,\tau_\Lambda} \mathbb{1}_{\tau_{\tilde{a}}^V < \tau_1^H}\right) \leq e^c e^{4\eta(1+\gamma)h/g}. \qquad (2.6.11)$$

for any $h > 1$, $\tilde{a} < \nu \leq g/\gamma$ and $\eta < \frac{1}{16}\left(g/(1+\gamma)\right)^2$. From (2.6.9) and (2.6.11), we obtain

$$\mathbb{E}_{(h,\nu)}\left(e^{\eta\,\tau_\Lambda}\right) \leq e^c e^{4\eta(1+\gamma)h/g}, \quad h > 1,\ \tilde{a} < \nu \leq g/\gamma,\ \eta < \frac{1}{16}\left(g/(1+\gamma)\right)^2. \qquad (2.6.12)$$

Lastly we consider the region to the right of Λ, where $\nu > \frac{g}{\gamma}$. When the velocity hits g/γ, either it is in Λ or $H > 1$ and the process has 'jumped over' the box's right edge. See Figure 2.1. Lemma .0.7 with $u = g/\gamma$ shows

$$\mathbb{E}_{(h,\nu)} \left(e^{\eta \tau_\Lambda} \mathbb{1}_{H_{\tau^V_{g/\gamma}} \leq 1} \right) \leq \mathbb{E}_{(h,\nu)} \, e^{\eta \, \tau^V_{g/\gamma}} \leq e^c \, e^{2 \frac{g}{\gamma} \nu}, \tag{2.6.13}$$

for $\eta < 2g^2/\gamma$. In addition, the strong Markov property at $\tau^V_{g/\gamma}$, along with (2.6.12), (2.6.13) and the Cauchy-Schwarz inequality imply,

$$\begin{aligned} \mathbb{E}_{(h,\nu)} \left(e^{\eta \tau_\Lambda} \mathbb{1}_{H_{\tau^V_{g/\gamma}} > 1} \right) &\leq e^c \, \mathbb{E}_{(h,\nu)} \left(e^{\eta \tau^V_{g/\gamma} + (4\eta(1+\gamma)/g) H_{\tau^V_{g/\gamma}}} \right) \\ \leq e^c \sqrt{\mathbb{E}_{(h,\nu)} \left(e^{2\eta \tau^V_{g/\gamma}} \right)} \sqrt{\mathbb{E}_{(h,\nu)} \, e^{(8\eta(1+\gamma)/g) H_{\tau^V_{g/\gamma}}}} &\leq e^c \, e^{2\nu \frac{g}{\gamma}} \sqrt{\mathbb{E}_{(h,\nu)} \, e^{(8\eta(1+\gamma)/g) H_{\tau^V_{g/\gamma}}}}, \end{aligned} \tag{2.6.14}$$

for $\eta < g^2/\gamma$ sufficiently small that the second term exists. We now define the allowable range of η and bound the second term in (2.6.14). Using (2.1.1), we have for all $(h, \nu) \in [0, \infty) \times (g/\gamma, \infty)$ and $t < \tau^V_{g/\gamma}$,

$$\begin{aligned} H_t + V_t/\gamma &= h + \nu/\gamma - B_t - tg/\gamma + (1 + 1/\gamma) L_t \\ &\leq h + \nu/\gamma + \sup_{t<\infty} \left(-B_t - tg/\gamma \right) + (1 + 1/\gamma) \sup_{t<\infty} \left(B_t - tg/\gamma \right). \end{aligned} \tag{2.6.15}$$

If η is small enough that $(1+1/\gamma) \times 16\eta(1+\gamma)/g < g/\gamma$, in other words $\eta < \frac{1}{16}(g/(1+\gamma))^2$, then since $E_1^* := \sup_{t<\infty} (-B_t - tg/\gamma)$ and $E_2^* := \sup_{t<\infty} (B_t - tg/\gamma)$ are exponentially distributed with mean $\gamma/2g$, (2.6.15) gives

$$\begin{aligned} \mathbb{E}_{(h,\nu)} \, e^{(8\eta(1+\gamma)/g)\left(\gamma^{-1} V_{t \wedge \tau^V_{g/\gamma}} + H_{t \wedge \tau^V_{g/\gamma}} \right)} &\leq e^{8\eta(1+\gamma)(\gamma h + \nu)/g\gamma} \mathbb{E} \, e^{(8\eta(1+\gamma)/g)(E_1^* + (1+1/\gamma)E_2^*)} \\ \leq e^{8\eta(1+\gamma)(\gamma h+\nu)/g\gamma} \sqrt{\mathbb{E} \, e^{(16\eta(1+\gamma)/g)E_1^*}} \sqrt{\mathbb{E} \, e^{(16\eta(1+\gamma)(1+1/\gamma)/g)E_2^*}} &\leq e^c e^{8\eta(1+\gamma)(\gamma h+\nu)/g\gamma}, \end{aligned} \tag{2.6.16}$$

for any $t > 0$, where the third bound above follows from Cauchy-Schwarz inequality. By Fatou's lemma,

$$\mathbb{E}_{(h,\nu)} \, e^{(8\eta(1+\gamma)/g) H_{\tau^V_{g/\gamma}}} \leq e^c e^{8\eta(1+\gamma)(\gamma h+\nu)/g\gamma}. \tag{2.6.17}$$

Combining (2.6.17) and (2.6.14),

$$\mathbb{E}_{(h,\nu)}\left(e^{\eta\tau_\Lambda}\mathbb{1}_{H_{\tau^V_{g/\gamma}}>1}\right)\le e^c\,e^{2\nu\frac{g}{\gamma}}\,e^{4\eta(1+\gamma)(\gamma h+\nu)/g\gamma},\tag{2.6.18}$$

for $h\ge 0,\nu>g/\gamma$ and $\eta<g^2/\gamma\wedge\frac{1}{16}(g/(1+\gamma))^2=\frac{1}{16}(g/(1+\gamma))^2$. From (2.6.13) and (2.6.18), we obtain

$$\mathbb{E}_{(h,\nu)}\left(e^{\eta\tau_\Lambda}\right)\le e^c\,e^{2\nu\frac{g}{\gamma}}\,e^{4\eta(1+\gamma)(\gamma h+\nu)/g\gamma},\quad h\ge 0,\ \nu>g/\gamma,\ \eta<\frac{1}{16}(g/(1+\gamma))^2.\tag{2.6.19}$$

Summarizing (2.6.7), (2.6.8), (2.6.12) and (2.6.19),

$$\mathbb{E}_{(h,\nu)}\,e^{\eta\tau_\Lambda}\le e^c\,e^{2\nu\frac{g}{\gamma}+4\eta(1+\gamma)(\gamma h+\nu)/g\gamma},\quad (h,\nu)\notin\Lambda,\ \eta<\frac{1}{16}(g/(1+\gamma))^2.\tag{2.6.20}$$

Notice that to introduce the $2\nu g/\gamma$ term in (2.6.20) when applying (2.6.7), (2.6.8), (2.6.12), we need only use the fact that $1=e^{-2\nu g/\gamma}e^{2\nu g/\gamma}<e^{2(g/\gamma)^2}e^{2\nu g/\gamma}$, since $\nu>-g/\gamma$.

We complete the proof of the theorem. By definition, $\tau_\Lambda(1)=1+\inf\{t\ge 0\,|\,(H_{1+t},V_{1+t})\in\Lambda\}$, so for any (h,ν) we have by (2.6.20),

$$\begin{aligned}&\mathbb{E}_{(h,\nu)}\,e^{\eta\tau_\Lambda(1)}=\mathbb{E}_{(h,\nu)}\left(\mathbb{1}_{(H_1,V_1)\in\Lambda}\,e^{\eta\tau_\Lambda(1)}\right)+\mathbb{E}_{(h,\nu)}\left(\mathbb{1}_{(H_1,V_1)\notin\Lambda}\,e^{\eta\tau_\Lambda(1)}\right)\\&\le e^\eta+e^\eta\mathbb{E}_{(h,\nu)}\left(\mathbb{1}_{(H_1,V_1)\notin\Lambda}\mathbb{E}_{(H_1,V_1)}\,e^{\eta\tau_\Lambda}\right)\le e^c\left(1+\mathbb{E}_{(h,\nu)}\,e^{2V_1\frac{g}{\gamma}+4\eta(1+\gamma)(\gamma H_1+V_1)/g\gamma}\right),\end{aligned}\tag{2.6.21}$$

for η sufficiently small. We use the crude bounds $V_t\le\nu+\sup_{s\le 1}\left(B_s+s\frac{g}{\gamma}\right)$, $L_t\le\sup_{s\le 1}\left(B_s+s\frac{g}{\gamma}\right)$ for $t\le 1$ and so $H_1\le h+\int_0^1 V_t\,dt-B_1+\sup_{s\le 1}\left(B_s+s\frac{g}{\gamma}\right)\le h+\nu-B_1+2\sup_{s\le 1}\left(B_s+s\frac{g}{\gamma}\right)$. From this, (2.6.21) and the finiteness of exponential moments of $-B_1$ and $\sup_{s\le 1}\left(B_s+s\frac{g}{\gamma}\right)$ we calculate,

$$\begin{aligned}\mathbb{E}_{(h,\nu)}\,e^{\eta\tau_\Lambda(1)}&\le e^c\left(1+e^{2\nu\frac{g}{\gamma}+4\eta(1+\gamma)(\gamma h+(1+\gamma)\nu)/g\gamma}\right)\\&\le e^c\left(1+e^{9\nu g/4\gamma+hg/4(1+\gamma)}\right):=F(h,\nu),\end{aligned}\tag{2.6.22}$$

for $\eta<\frac{1}{16}(g/(1+\gamma))^2$ and any $(h,\nu)\in S$. □

The last ingredient needed in the proof of exponential ergodicity is a *minorization condition* (e.g. Assumption 2 of Hairer and Mattingly (2011), or Down et al. (1995) for a different formulation): There exist $t_0, \alpha > 0$ and a non-trivial measure μ such that

$$P^{t_0}((h,\nu),\,\cdot\,)) \geq \alpha\mu(\cdot), \qquad (h,\nu) \in \Lambda. \tag{2.6.23}$$

In essence, (2.6.23) ensures two versions of the process started within Λ can be coupled. Typically, (2.6.23) is checked by showing P^t has a continuous density with respect to Lebesgue measure for each $t > 0$. As stated before, these techniques are not available in our setup. Instead Lemmas 2.6.3 and 2.6.2, which use the renewal structure of our process in a crucial way, establish the following stronger version of (2.6.23).

Lemma 2.6.5. *Fix Λ as in Lemma 2.6.3. There exists a $t_0 > 0$ a non-trivial measure $\tilde{\mu}$ such that for all measurable $A \subset S$,*

$$P^t((h,\nu),\, A) \geq \tilde{\mu}(A) \qquad (h,\nu) \in \Lambda,\, t \geq t_0.$$

Proof. Theorem 2.6.2 shows that the process $(H_t, V_t)_{t\geq 0}$ is ergodic which, via Theorem 6.1 of Meyn and Tweedie (1993a), implies that there is a skeleton chain that is irreducible. Moreover, from Lemma 2.6.3 (c), Λ is 'petite.' The Lemma now follows from Proposition 6.1 of Meyn and Tweedie (1993a) upon noting that our process in positive Harris recurrent, which was proved in Lemma 2.6.3 (d). □

Proof of Theorem 2.1.6. The theorem follows from Theorems 5.2, 6.2 of Down et al. (1995), which extend the discrete time drift and minorization conditions of Harris's theorem to continuous-time processes and show how these can be used to obtain exponential ergodicity (see also Hairer and Mattingly (2011) for a discrete-time formulation). Specifically, Lemma 2.6.3 (c) shows that Λ is 'petite'. Theorem 6.2 of Down et al. (1995) (with $f \equiv 1$ and $\delta = 1$) and Lemma 2.6.4 show the function $G(h,\nu) = 1 + \frac{1}{\eta}\left(\mathbb{E}_{(h,\nu)}\, e^{\eta\tau_\Lambda(1)} - 1\right)$ is a Lyapunov function for the process which satisfies the following *drift condition* for every $t_0 > 0$,

$$P^t G \leq \lambda_{t_0} G + c\mathbb{1}_\Lambda \quad t \leq t_0,\; c > 0, \lambda_{t_0} < 1. \tag{2.6.24}$$

The result then follows from Theorem 5.2 of Down et al. (1995). We note here that the referenced theorem requires an 'aperiodicity' type condition for the semigroup P^t, given on p. 1675 of the reference (which is not the same as the more standard notion of aperiodicity defined in (2.6.26) below). However, the first line in the proof of Theorem 5.2 of Down et al. (1995) makes clear that the sole use of the 'aperiodicity' condition is to ensure there exists a $t_0 > 0$ such that the discrete-time process with transitions $\{P^{kt_0}\}_{k\geq 1}$ is geometrically ergodic. In other words,

$$\|P^{kt_0}((h,\nu),\ \cdot) - \pi\|_{TV} \leq G(h,\nu)D_0 r^k, \quad k \geq 1,\ (h,\nu) \in S, \tag{2.6.25}$$

for some $t_0 > 0$ and constants $D_0 \in (0,\infty)$ and $r \in (0,1)$, with G as in (2.6.24).

For clarity, we verify here that (2.6.25) holds, from which the proof of Theorem 5.2 in Down et al. (1995) can proceed as written. First, we recall the definition of aperiodicity for a discrete-time Markov chain (see Section 5.4.3 of Meyn and Tweedie (2009), Theorem 5.4.4 and the discussion preceding it): A process with transitions $\{Q^k\}_{k\geq 1}$ is called *aperiodic* if there exists a set C such that (2.6.23) holds for Q^k in place of P^{t_0} for some $k \geq 1$ and C in place of Λ, and such that,

$$gcd\{n \geq 1 \,:\, Q^n((h,\nu),\cdot\,) \geq \alpha_n \mu(\cdot), \quad (h,\nu) \in C \quad \text{some } \alpha_n > 0\} = 1. \tag{2.6.26}$$

Lemma 2.6.5 shows (2.6.26) is satisfied for the chain with transition laws $Q^k = P^{kt_0}$, $C = \Lambda$ and $\alpha_n = 1$ for all n, where Λ and t_0 are as given in the Lemma. Thus, the discretely sampled chain with transition laws $\{P^{kt_0}\}_{k\geq 1}$ is aperiodic. (2.6.25) follows immediately from this observation in conjunction with (2.6.24) and Theorem 16.0.1 (ii), (iv) of Meyn and Tweedie (2009). The proof of Theorem 5.2 in Down et al. (1995) now proceeds exactly as written, giving the result. $\square$

CHAPTER 3

Dimension-free and local convergence of reflected Brownian motion

3.1 Introduction

We say a continuous stochastic process X is a solution to $\mathrm{RBM}(\Sigma, \mu, R)$ if it satisfies

$$X(x,t) = x + \mu t + DB(t) + RL(x,t) \tag{3.1.1}$$

for each $t > 0$ and $x \in \mathbb{R}^d_+ := \{x \in \mathbb{R}^d \mid x_i \geq 0,\ i = 1 \dots d\}$. Here $\mu \in \mathbb{R}^d, D, R \in \mathbb{R}^{d\times d}$, B is a d-dimensional Brownian motion and $\Sigma = DD^T$ is positive definite. We assume that $R = I - P^T$ for a matrix P that is sub-stochastic (i.e. non-negative entries and row sums are bounded above by one) and transient (i.e. $P^n \to 0$ as $n \to \infty$). L is the local time constraining X to the positive orthant $\mathbb{R}^d_+$: It is a non-decreasing, continuous process adapted to the natural filtration of the Brownian motion B such that

$$L(0) = 0, \qquad \int_0^t X_i(x,s) dL_i(s) = 0 \text{ for all } t > 0, 1 \leq i \leq d.$$

RBMs of the form (3.1.1) arise in a variety of situations, including heavy-traffic limits of queue-length processes in generalized Jackson networks with d servers Reiman, O. (1984); Harrison and Williams (1987a), and gaps between $d+1$ competing particles in rank-based diffusions (e.g. Karatzas, I. and Pal, S. and Shkolnikov, M. (2016); Sarantsev, A. (2019)).

In this chapter, we are interested in the effect of dimension on convergence rates to stationarity for reflected Brownian motions (RBMs) from a variety of initial configurations. This is a natural consideration for steady state sampling and evaluating steady state performance for high dimensional RBMs. Towards this end, we will implicitly consider a family of processes $X^{(d)} \sim \mathrm{RBM}(\Sigma^{(d)}, \mu^{(d)}, R^{(d)})$ indexed by the dimension $d \geq 1$. For notational convenience, we will suppress the suffix (d) in further discussion.

3.1.1 Convergence rates for RBM: work to date

There is a large literature studying diffusions with oblique reflections, in cases both more specific and more general than (3.1.1), and we give only a brief background describing previous work most relevant to the current article. Harrison and Reiman (1981) first proved (3.1.1) has a unique strong solution. Harrison and Williams (1987a, Section 6) showed (3.1.1) has a stationary distribution if and only if $R^{-1}\mu < 0$, and in that case the stationary distribution is unique.

To study convergence rates of X to its stationary distribution, one can apply general methods like Harris' theorem via using appropriate Lyapunov functions and minorization conditions Meyn and Tweedie (2012). For example, Budhiraja and Lee (2007) uses this methodology to give exponentially fast convergence of $X(x, \cdot)$ to the stationary random variable in a weighted total variation norm starting from any $x \in \mathbb{R}_d^+$. However, the rate of convergence is not explicit, as is typical for such methods, and in particular has unknown dimension dependence. See also Sarantsev, A. (2017b) for a similar treatment.

In Blanchet and Chen (2020), the authors obtained explicit dimension dependent convergence rates to stationarity in L^1-Wasserstein distance when the RBM satisfies 'uniformity conditions in dimension' on the model parameters Σ, μ, R (discussed here in more detail in Example 3.2.2). Their key insight was to consider *synchronous couplings* of the RBM X (i.e. driven by the same Brownian motion) started from distinct points $x, y \in \mathbb{R}_+^d$, with $x \leq y$ (coordinate wise ordering). They used the fact that synchronous couplings preserve ordering in time, that is, $X(x,t) \leq X(y,t)$ for all $t \geq 0$. Moreover, there are contractions in L^1 distance between the synchronously coupled processes (under their uniformity assumptions) when the dominating process $X(y, \cdot)$ has hit all faces of the orthant $\mathbb{R}_+^d$. Building on this idea, Banerjee and Budhiraja (2020) used a weighted Lyapunov function and excursion theoretic control of the synchronously coupled processes to give convergence rates in L^1-Wasserstein distance for the general process (3.1.1) which depend explicitly on μ, R, Σ, d. In particular, this approach greatly improved the rates for the models considered in Blanchet and Chen (2020) from polynomial in d to logarithmic in d.

3.1.2 Dimension-free local convergence for RBM

Typically, growing dimension slows down the rate of convergence for the whole system, as reflected in the bounds obtained in Blanchet and Chen (2020); Banerjee and Budhiraja (2020), but one might observe a much faster convergence rate to equilibrium of *local statistics* of the system. In Section 3.2, we describe and investigate a class of RBMs for which convergence rates of local statistics do not depend on the underlying dimension of the entire system. We call this phenomenon dimension-free local convergence.

Mathematically, this is challenging as the local evolution is no longer Markovian and the techniques in Blanchet and Chen (2020); Banerjee and Budhiraja (2020) cannot be readily applied. We make a crucial observation that certain weighted L^1 distances (see $\|\cdot\|_{1,\beta}$ defined in Section 3.2.1) between synchronously coupled RBMs show dimension-free contraction rates. The evolution of such weighted distances are tracked in time for synchronously coupled RBMs $X(0,\cdot)$ and $X(x,\cdot)$ for $x \in \mathbb{R}^d_+$. It is shown that for this distance to decrease by a dimension-free factor of its original value, only a subset of co-ordinates of $X(x,\cdot)$, whose cardinality depends on the value of the original distance, need to hit zero. This is in contrast with the unweighted L^1 distance considered in Blanchet and Chen (2020); Banerjee and Budhiraja (2020) where all the coordinates need to hit zero to achieve such a contraction, thereby slowing down the convergence rate. Consequently, by tracking the hitting times to zero of a time dependent number of co-ordinates, one achieves dimension-free convergence rates in this weighted L^1 distance as stated in Theorem 3.2.1. This, in turn, gives dimension-free local convergence as is made precise in (3.2.4). In Section 3.2.4, Theorem 3.2.1 is applied to two important classes of RBM to obtain explicit convergence rates.

3.1.3 Perturbations from stationarity for the Symmetric Atlas Model

As a first step in studying dimension-free convergence rates for RBMs which do not satisfy the assumptions of Section 3.2, we focus attention in Section 3.3 on the Symmetric Atlas model. This is a rank-based diffusion comprising $d+1$ Brownian particles where the least ranked particle performs a Brownian motion with constant positive drift and the remaining particles perform standard Brownian motions. The gaps between the ordered particles collectively evolve as a RBM which converges in total variation distance to an explicit stationary measure (3.3.3) Pal and Pitman (2008). Interestingly, the gap process of the infinite-dimensional version of the Symmetric Atlas model obtained in Pal and Pitman (2008) has infinitely many stationary measures Sarantsev and Tsai (2017), only one of which is a weak limit of the stationary measure (3.3.3) of the d-dimensional system as $d \to \infty$. This leads to the heuristic that, for large d, the d-dimensional gap process with initial distribution 'close' to the projection (onto the first d co-ordinates) of one of the other infinite-dimensional stationary measures spends a long time near this projection before converging to (3.3.3). From this heuristic, one expects that dimension-free convergence rates for associated statistics can only be obtained if the initial gap distribution is 'close' to the stationary measure (3.3.3) in a certain sense. Evidence for this heuristic is provided in the few available results on 'uniform in dimension' convergence rates of some rank-based diffusions Jourdain and Malrieu (2008); Jourdain and Reygner (2013). In both these papers, under strong convexity assumptions on the drifts of the particles, dimension-free exponential ergodicity was proven for the joint density of the particle system when the initial distribution is close to the stationary distribution as quantified by the Dirichlet energy functional (see Jourdain and Malrieu (2008, Theorem 2.12) and Jourdain and Reygner (2013, Corollary 3.8)). The Symmetric Atlas model lacks such convexity in drift and hence, the dimension-free Poincaré inequality for the stationary density, that is crucial to the methods of Jourdain and Malrieu (2008); Jourdain and Reygner (2013), does not apply. We take a very different approach which involves analyzing the long term behavior of pathwise derivatives of the RBM in initial conditions. Using this analysis, we obtain polynomial convergence rates to stationarity in L^1-Wasserstein distance when the initial distribution of the gaps between particles is in an appropriate perturbation class (defined in Definition 3.3.1) of the stationary measure. Although we do not yet have lower bounds on convergence rates, we strongly believe that the optimal rates are indeed polynomially decaying in time (see Remark 3.3.2).

We mention here that Blanchet et al. (2020) has recently used the derivative process to study convergence rates for RBMs satisfying strong uniformity conditions in dimension (which do not hold for the Symmetric Atlas model). Our analysis of the derivative is based on a novel connection with a random walk in a random environment generated by the times and locations where the RBM hits faces of $\mathbb{R}_+^d$ (see Section 3.3.2). We believe our analysis can be combined with that of Blanchet et al. (2020) to study ergodicity properties of more general classes of RBM. This is deferred to future work.

We also mention the work of Pal and Sarantsev (2019) who obtained a dimension-free Talagrand type transportation cost-information inequality for reflected Brownian motions. Such inequalities, however, are more useful in dimension-free concentration of measure phenomena as opposed to dimension-free rates of convergence to stationarity.

3.1.4 Chapter generic notation

Here we list notation for general concepts and conventions. Inequalities for vectors are evaluated element-wise. For a square matrix A, $A|_k$ is the $k \times k$ northwest quadrant. For a vector v, $v|_k$ is the projection of v onto the first k coordinates. Other conventions include $x \vee y = \max(x, y)$, $x \wedge y = \min(x, y)$, $\lfloor x \rfloor = \max\{k \in \mathbb{Z} \mid k \leq x\}$ and $x^+ = \max(0, x)$.

For $x \in \mathbb{R}^k$, we write the supremum norm as $\|x\|_\infty = \max_{1 \leq i \leq k} |x_i|$ and the ℓ^1 norm as $\|x\|_1 := \sum_{i=1}^k |x_i|$. For a fixed $\beta \in (0, 1)$, define a weighted ℓ^1 norm by $\|x\|_{1,\beta} = \sum_{i=1}^k \beta^i |x_i|$ and weighted supremum norm by $\|x\|_{\infty,\beta} = \max_{1 \leq i \leq k} \beta^i |x_i|$.

For X a RBM(Σ, μ, R) started at $x \in \mathbb{R}_+^d$ and any $k \in \{1 \ldots d\}$, we write $X(\infty)$ for the random variable with the stationary distribution. Write $X|_k(\cdot, x)$ for the process restricted to its first k coordinates.

3.2 Dimension-free local convergence rates for RBM

3.2.1 A weighted norm governing dimension-free convergence

Our investigation of dimension-free convergence relies on the analysis of the weighted distance $\|X(x, \cdot) - X(X(\infty), \cdot)\|_{1,\beta}$ in time, for appropriate choices of $\beta \in (0, 1)$. Towards this end, we will analyze the following functionals:

$$u_\beta(x, t) = \|R^{-1}\left(X(x, t) - X(0, t)\right)\|_{1,\beta} := \sum_{i=1}^d \beta^i \left| \left[R^{-1}\left(X(x, t) - X(0, t)\right)\right]_i \right|, \tag{3.2.1}$$

$$u_{\pi,\beta}(t) = u_\beta(X(\infty), t), \qquad t \geq 0. \tag{3.2.2}$$

In the following, when β is clear from context, we will suppress dependence on β and write u for u_β and u_π for $u_{\pi,\beta}$. The above functionals are convenient because the vector $R^{-1}(X(x,t)-X(0,t))$ is co-ordinate wise non-negative and non-increasing in time (see Theorem 3.4.1 (iii)). This fact and the triangle inequality can be used to show for any $x \in \mathbb{R}_+^d, t \geq 0$ (see (3.4.40)),

$$\|(X(x,t)-X(X(\infty),t))\|_{1,\beta} \leq u(x,t)+u_\pi(t).$$

We are interested in conditions under which there exists a d-independent $\beta \in (0,1)$ and a function $f:\mathbb{R}_+ \mapsto \mathbb{R}_+$ not depending on the dimension d of X such that $f(t) \to 0$ as $t \to \infty$ and

$$\mathbb{E}\left[\|(X(x,t)-X(X(\infty),t))\|_{1,\beta}\right] \leq Cf(t), \qquad t \geq t_0, \tag{3.2.3}$$

where $C, t_0 \in (0,\infty)$ are constants not depending on d. This, in particular, gives dimension-free local convergence in the following sense: For any $k \in \{1,\ldots,d\}$, consider any function $\phi:\mathbb{R}_+^k \mapsto [0,\infty)$ which is L^1-Lipschitz, i.e., there exists $L_\phi > 0$ such that

$$|\phi(x)-\phi(y)| \leq L_\phi \|x-y\|_1, \qquad x,y \in \mathbb{R}_+^k.$$

Recall that the L^1-Wasserstein distance between two probability measures μ and ν on $\mathbb{R}_+^k$ is given by

$$W_1(\mu,\nu) = \inf\left\{\int_{\mathbb{R}_+^k \times \mathbb{R}_+^k} \|\mathbf{x}-\mathbf{y}\|_1 \gamma(\mathrm{d}\,\mathbf{x},\mathrm{d}\,\mathbf{y}) \ : \gamma \text{ is a coupling of } \mu \text{ and } \nu\right\}.$$

Denote the law of a random variable Z by $\mathcal{L}[Z]$. Then, (3.2.3) implies

$$\begin{aligned} &W_1\left(\mathcal{L}[\phi(X|_k(x,t))], \mathcal{L}[\phi(X(\infty))]\right) \\ &\qquad \leq \mathbb{E}\left[|\phi(X|_k(x,t))-\phi(X|_k(X(\infty),t))|\right] \leq C\beta^{-k}L_\phi f(t), \qquad t \geq t_0. \end{aligned} \tag{3.2.4}$$

3.2.2 Parameters and Assumptions

We now define the parameters that govern dimension-free local convergence which, in turn, are defined in terms of the original model parameters (Σ, μ, R) of the associated RBM. We abbreviate $\sigma_i = \sqrt{\Sigma_{ii}}, i = 1, \ldots, d$. Define for $1 \leq k \leq d$,

$$
\begin{aligned}
b^{(k)} &:= -\left(R|_k\right)^{-1} \mu|_k, \qquad b = b^{(d)}, \\
\underline{b}^{(k)} &:= \min_{1\leq i\leq k} b_i^{(k)}, \qquad a^{(k)} := \max_{1\leq i\leq k} \frac{1}{b_i^{(k)}} \sum_{j=1}^{k} (R|_k)^{-1}_{ij}\, \sigma_j.
\end{aligned}
\tag{3.2.5}
$$

To get a sense of why these parameters are crucial, recall that our underlying strategy is to obtain contraction rates of $u(x, \cdot)$ defined in (3.2.1) by estimating the number of times a subset of the co-ordinates of $X(x, \cdot)$, say $\{X_1(x, \cdot), \ldots, X_k(x, \cdot)\}, k \leq d$, hit zero. However, this subset does not evolve in a Markovian way. Thus, we use monotonicity properties of RBMs to couple this subset with a $\mathbb{R}^k_+$-valued reflected Brownian motion $\bar{X}(x|_k, \cdot)$, started from $x|_k$ and defined in terms of $\mu|_k, D|_k, R|_k$ and (a possible restriction of) the same Brownian motion driving $X(x, \cdot)$, such that $X_i(x, t) \leq \bar{X}_i(x|_k, t)$ for all $1 \leq i \leq k$ (see Theorem 3.4.2). The analysis in Banerjee and Budhiraja (2020) shows that the parameters defined in (3.2.5) with $k = d$ can be used to precisely estimate the minimum number of times all co-ordinates of $X(x, \cdot)$ hit zero by time t as t grows. Thus, for any $1 \leq k \leq d$, the parameters (3.2.5) can be used to quantify analogous hitting times for the process $\bar{X}(x|_k, \cdot)$ which, by the above coupling, gives control over corresponding hitting times of $\{X_1(x, \cdot), \ldots, X_k(x, \cdot)\}$.

We list below two sets of assumptions on the model parameters (Σ, μ, R) which guarantee dimension-free local convergence.

Assumption 3.2.1. *There exist d-independent constants $\underline{\sigma}, \overline{\sigma}, b_0 > 0$, $r^* \geq 0$, $M, C \geq 1$, $k_0 \in \{2, \ldots, d\}$ and $\alpha \in (0, 1)$ such that for all $d \geq k_0$,*

I. $R^{-1}_{ij} \leq C\alpha^{j-i}$ *for* $1 \leq i \leq j \leq d$,

II. $R^{-1}_{ij} \leq M$ *for* $1 \leq i, j \leq d$,

III. $\underline{b}^{(k)} \geq b_0 k^{-r^*}$ *for* $k = k_0, \ldots, d$,

IV. $\sigma_i \in [\underline{\sigma}, \overline{\sigma}]$ *for* $1 \leq i \leq d$.

We explain why Assumption 3.2.1 is 'natural' in obtaining dimension-free local convergence. Since P is a transient and substochastic, it can be associated to a killed Markov chain on $\{0\} \cup \{1, \ldots, d\}$ with transition matrix P on $\{1, \ldots, d\}$ and killed at 0 (i.e. probability of going from state $k \in \{1, \ldots, d\}$ to 0 is $1 - \sum_{l=1}^{d} P_{kl}$ and $P_{00} = 1$). Moreover, since P is transient and $R = I - P^T$, we have $R^{-1} = \sum_{n=0}^{\infty} (P^T)^n$. This representation shows that R^{-1}_{ij} is the expected number of visits to site i starting from j of this killed Markov chain. For fixed $x \in \mathbb{R}^d_+$ and $k << d$, consider a local statistic of the form $\phi(X|_k(x,t))$ as in (3.2.4). For this statistic to stabilize faster than the whole system, we expect the influence of the far away co-ordinates $X|_j(x, \cdot), j >> k$, to diminish in an appropriate sense as k increases. This influence is primarily manifested through the oblique reflection arising out of the R matrix in (3.1.1). I of Assumption 3.2.1 quantifies this intuition by requiring that the expected number of visits to state i starting from state $j > i$ of the associated killed Markov chain decreases geometrically with $j - i$. This is the case, for example, when this Markov chain started from $j > i$ has a uniform 'drift' away from i towards the cemetery state. See Example 3.2.1. In more general cases, one can employ Lyapunov function type arguments Meyn and Tweedie (2012) to the underlying Markov chain to check I.

II above implies that the killed Markov chain starting from state j spends at most M expected time at any other site $i \in \{1, \ldots, d\}$ before it is absorbed in the cemetery state 0. This expected time, as our calculations show, is intimately tied to decay rates of $\| (X(x, \cdot) - X(X(\infty), \cdot)) \|_{1,\beta}$.

As noted in Harrison and Williams (1987a); Blanchet and Chen (2020); Banerjee and Budhiraja (2020), the 'renormalized drift' vector b characterizes positive recurrence of the whole system. Through III above, we allow for a power law type co-ordinate wise lower bound of the renormalized drift vector $b^{(k)}$ of the projected system $X|_k(x, \cdot)$ as k grows. In particular, if $\underline{b}^{(k)}$ is uniformly lower bounded by b_0, we can take $r^* = 0$.

IV above is a quantitative 'uniform ellipticity' condition on the co-ordinates of the driving noise $DB(\cdot)$.

Note that we do not need to make any assumptions on the correlations of the driving noise, i.e. on $\sigma_{ij}/(\sigma_i, \sigma_j)$ for $i < j$. This can be understood upon inspection of our proof technique where the drift and the reflection 'overpower' the diffusivity in long time contraction properties of $\| (X(x,\cdot) - X(X(\infty),\cdot)) \|_{1,\beta}$. The following assumption is a strengthening of Assumption 3.2.1 which, when satisfied, will lead to significantly better convergence rates to stationarity.

Assumption 3.2.2. *Suppose Assumption 3.2.1 holds. In addition assume M, which does not depend on d, may be chosen large enough that*

II'. $\max_{1\le i\le d} \sum_{j=1}^{d} R_{ij}^{-1} \le M$.

This is satisfied, for example, when there exist positive d-independent constants j_0, p_0 such that the underlying killed Markov chain has jump size bounded by j_0 at each step, and a probability of at least p_0 of reaching 0 in one step from any starting site in $\{1,\dots,d\}$. See Example 3.2.2 for such a RBM.

3.2.3 Main results

We present here our first main result exhibiting explicit dimension-free convergence rates in the weighted distance $\|\cdot\|_{\sqrt{\alpha}}$ (α defined in I of Assumption 3.2.1) for RBMs satisfying Assumption 3.2.1 or 3.2.2. We first define some constants that will appear in Theorem 3.2.1. They are needed to bound moments of weighted norms of the stationary random variable $X(\infty)$ and are derived in Lemma 3.4.8.

Suppose Assumption 1 holds, with $k_0 \in \{2\dots d\}$ and $\alpha \in (0,1)$ defined therein. Set

$$L_1 := k_0^{r^*+1} + \sum_{i=k_0}^{d} i^{3+r^*} \alpha^{i/8}. \tag{3.2.6}$$

If in addition Assumption 3.2.2 holds, define

$$L_2 := k_0^{r^*} + \sum_{i=k_0}^{d} i^{2+r^*} \alpha^{i/8}. \tag{3.2.7}$$

Also, for $B \in (0,\infty)$, define the set

$$\mathcal{S}(b,B) := \left\{ x \in \mathbb{R}_d^+ : \sup_{1\le i\le d} \underline{b}^{(i)} \|x|_i\|_\infty \le B \right\}. \tag{3.2.8}$$

Theorem 3.2.1 directly implies dimension-free local convergence rates as given by (3.2.4).

Theorem 3.2.1. *Suppose Assumption 3.2.1 holds for X, an $\mathrm{RBM}(\Sigma, \mu, R)$, with $\alpha \in (0,1)$ defined therein. Recall the weighted distance $\|\cdot\|_{1,\sqrt{\alpha}}$ (taking $\beta = \sqrt{\alpha}$) defined in Section 3.1.4.*

Fix any $B \in (0,\infty)$. Then there exist constants $C_0, C_0', C_1 > 0$ not depending on d, r^* or B such that with $t_0' = t_0'(r^*) = C_0'(1+r^*)^{8+4r^*}$ and $L_1, L_2, \mathcal{S}(b,B)$ as defined in (3.2.6)-(3.2.8), we have for any $x \in \mathcal{S}(b,B)$ and any $d > t_0'^{1/(4+2r^*)}$,

$$\mathbb{E}\left[\|X(x,t) - X(X(\infty),t)\|_{1,\sqrt{\alpha}}\right] \leq \begin{cases} C_1\left(L_1\sqrt{1+t^{1/(4+2r^*)}} + \|x\|_\infty \exp\left\{B/\underline{\sigma}^2\right\}\right) \exp\{-C_0 t^{1/(4+2r^*)}\}, & t_0' \leq t < d^{4+2r^*}, \\ \\ C_1\left(L_1\sqrt{1+t^{1/(4+2r^*)}} + \|x\|_\infty \exp\left\{B/\underline{\sigma}^2\right\}\right) \exp\left\{-C_0 \frac{t}{d^{3+2r^*}}\right\}, & t \geq d^{4+2r^*}. \end{cases} \tag{3.2.9}$$

If instead Assumption 3.2.2 holds, with $t_1' = t_1'(r^*) = C_0'(1+r^*)^{2+4r^*}$, we have for $d > t_1'^{1/(1+2r^*)}$,

$$\mathbb{E}\left[\|X(x,t) - X(X(\infty),t)\|_{1,\sqrt{\alpha}}\right] \leq \begin{cases} C_1\left(L_2\sqrt{1+t^{1/(1+2r^*)}} + \|x\|_\infty \exp\left\{B/\underline{\sigma}^2\right\}\right) \exp\left\{-C_0 \frac{t^{1/(1+2r^*)}}{\log t}\right\}, & t_1' \leq t < d^{1+2r^*}, \\ \\ C_1\left(L_2\sqrt{1+t^{1/(1+2r^*)}} + \|x\|_\infty \exp\left\{B/\underline{\sigma}^2\right\}\right) \exp\left\{-C_0 \frac{t}{d^{2r^*}\log d}\right\}, & t \geq d^{1+2r^*}. \end{cases} \tag{3.2.10}$$

Remark 3.2.1. *Bounds analogous to those in Theorem 3.2.1 hold using the norm $\|\cdot\|_{1,\beta}$ for any $\beta \in (\alpha, 1)$, with appropriately adjusted constants depending on β, and the choice $\beta = \sqrt{\alpha}$ is merely for simplicity of exposition. In fact, our proofs are in terms of two parameters $\beta \in (\alpha,1)$ and $\delta \in (\beta, 1)$, which can be appropriately chosen for the specific RBM under consideration to optimize the obtained bounds.*

3.2.4 Applications of Theorem 3.2.1

Here, we present two examples of RBMs that arise in diverse applications, where we can apply Theorem 3.2.1 to obtain explicit dimension-free convergence rates.

Example 3.2.1 (Asymmetric Atlas model)**.** We consider Atlas-type models, which are interacting particle systems represented by the following SDE:

$$Z_k(t) = Z_k(0) + \mathbb{1}_{k=1}t + B_k^*(t) + pL_{(k-1,k)}(t) - qL_{(k,k+1)}(t), \quad t \geq 0, \tag{3.2.11}$$

for $1 \leq k \leq d+1$, $p \in (0,1), q = 1-p$. Here, $L_{(0,1)}(\cdot) \equiv L_{(d+1,d+2)}(\cdot) \equiv 0$, and for $1 \leq k \leq d$, $L_{(k,k+1)}(\cdot)$ is a continuous, non-decreasing, adapted process that denotes the collision local time between the k-th and $(k+1)$-th co-ordinate processes of Z, namely $L_{(k,k+1)}(0) = 0$ and $L_{(k,k+1)}(\cdot)$ can increase only when $Z_k = Z_{k+1}$. $B_k^*(\cdot)$, $1 \leq k \leq d+1$, are mutually independent standard one dimensional Brownian motions. Each of the $d+1$ ranked particles with trajectories given by $(Z_1(\cdot), \ldots, Z_{d+1}(\cdot))$ evolves as an independent Brownian motion (with the particle 1 having unit positive drift) when it is away from its neighboring particles, and interacts with its neighbors through possibly asymmetric collisions. The Symmetric Atlas model, namely the case $p = 1/2$, was introduced in Fernholz (2002) as a mathematical model for stochastic portfolio theory. The Asymmetric Atlas model, namely the case $p \in (1/2, 1)$, was introduced in Karatzas, I. and Pal, S. and Shkolnikov, M. (2016). It was shown that it arises as scaling limits of numerous well known interacting particle systems involving asymmetrically colliding random walks Karatzas, I. and Pal, S. and Shkolnikov, M. (2016, Section 3). Since then, this model has been extensively analyzed: see Karatzas, I. and Pal, S. and Shkolnikov, M. (2016); Ichiba et al. (2013b,a); Sarantsev, A. (2017a) and references therein.

The gaps between the particles, defined by $X_i(\cdot) = Z_{i+1}(\cdot) - Z_i(\cdot), 1 \leq i \leq d$, evolve as an RBM$(\Sigma, \mu, R)$ with Σ given by $\Sigma_{ii} = 2$ for $i = 1 \ldots d$, $\Sigma_{ij} = -1$ if $|i-j| = 1$, $\Sigma_{ij} = 0$ if

$|i-j|>1$, μ given by $\mu_1=-1, \mu_j=0$ for $j=2,\ldots,d$, and $R=I-P^T$, where

$$P_{ij}=\begin{cases} p & j=i+1, \\ 1-p & j=i-1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.2.12}$$

In this chapter, we are interested in the ergodicity of the gap process X. In the current example, we study the Asymmetric Atlas model. The Symmetric Atlas model is treated in Section 3.3.

Recall that the reflection matrix $R=I-P^T$ is associated with a killed Markov chain. For the Asymmetric Atlas model, this Markov chain has a more natural description as a random walk on $\{0,1,\ldots,d+1\}$ which increases by one at each step with probability p and decreases by one with probability $1-p$, and is killed when it hits either 0 or $d+1$. Then for $1\le i,j\le d$, R^{-1}_{ij} is the expected number of visits to i starting from j by this random walk before it hits 0 or $d+1$. Since $p>1-p$, the random walk has a drift towards $d+1$, which suggests I, II of Assumption 3.2.1 hold. This is confirmed by direct computation, which gives for $q=1-p$,

$$R^{-1}_{ij}=\begin{cases} \frac{(q/p)^{j-i}}{p-q}\frac{\left(1-(q/p)^i\right)\left(1-(q/p)^{d+1-j}\right)}{1-(q/p)^{d+1}} \le \frac{(q/p)^{j-i}}{p-q} & 1\le i\le j\le d, \\ \frac{(p/q)^{i-j}}{p-q}\frac{\left((p/q)^j-1\right)\left((p/q)^{d+1-i}-1\right)}{(p/q)^{d+1}-1} \le \frac{1}{p-q} & 1\le j<i\le d. \end{cases} \tag{3.2.13}$$

Now I, II and IV of Assumption 3.2.1 holds with $M=C=\frac{1}{p-q}$, $\alpha=\frac{q}{p}$ and $\underline{\sigma}=\overline{\sigma}=\sqrt{2}$. Furthermore, the restriction $P|_k$ is defined exactly as in (3.2.12) with k in place of d. Thus $(R|_k)^{-1}$ is given by (3.2.13) with k in place of d, and $b^{(k)}=-\left(R|_k\right)^{-1}\mu|_k$ is the first column of $(R|_k)^{-1}$. This entails,

$$\begin{aligned} b^{(k)}_i &= \frac{(p/q)^{i-1}}{p-q}\frac{\left((p/q)-1\right)\left((p/q)^{k+1-i}-1\right)}{(p/q)^{k+1}-1} \ge \frac{1}{q}\left(\frac{p}{q}\right)^{k-1}\frac{\left((p/q)-1\right)}{(p/q)^{k+1}} \\ &= \frac{p-q}{p^2} =: b_0>0, \qquad 1\le i\le k,\ 1\le k\le d. \end{aligned} \tag{3.2.14}$$

Thus $\underline{b}^{(k)}\ge b_0$ for all $1\le k\le d$, uniformly in d. This shows that III of Assumption 3.2.1 holds with b_0 specified by (3.2.14) and $r^*=0$. Moreover, it follows from the first equality in (3.2.14) that $b_i^{(k)} \le p/(p-q)$ for all $1 \le k \le d$ and $1 \le i \le k$. Therefore, recalling the definition of $\mathcal{S}(b,\cdot)$ from (3.2.8), for any $x \in \mathbb{R}_+^d$,

$$x \in \mathcal{S}(b, p\|x\|_\infty/(2p-1)).$$

Finally we note Assumption 3.2.2 does not hold here. It can be checked from (3.2.13) that $\sum_{j=1}^d R_{ij}^{-1}$ grows linearly in i and hence, the row sums of R^{-1} are not uniformly bounded by a dimension-independent constant. This stands in contrast with Example 3.2.2.

The above observations result in the next theorem, which follows directly from Theorem 3.2.1.

Theorem 3.2.2. *Suppose X is the RBM for the asymmetric Atlas model. Then there exist constants $\bar{C}, \bar{C}_0, t_0' > 0$ depending on p but not on d such that for $d > t_0'$,*

$$\mathbb{E}\left[\|X(x,t) - X(X(\infty),t)\|_{1,\sqrt{\frac{1-p}{p}}}\right] \le \begin{cases} \bar{C}\left(\sqrt{1+t^{1/4}} + \|x\|_\infty e^{p\|x\|_\infty/(4p-2)}\right) e^{-\bar{C}_0 t^{1/4}}, & t_0' \le t < d^4, \\ \\ \bar{C}\left(\sqrt{1+t^{1/4}} + \|x\|_\infty e^{p\|x\|_\infty/(4p-2)}\right) e^{-\bar{C}_0 t/d^3}, & t \ge d^4. \end{cases} \tag{3.2.15}$$

Example 3.2.2 (Blanchet-Chen type conditions)**.** Here we consider RBM(Σ,μ,R) with the system parameters satisfying certain 'uniformity' assumptions in dimension similar to those of Blanchet and Chen (2020). In addition, we assume P is a 'band matrix' (see assumption a) below).

With the notation of Assumption 3.2.1: Suppose there exist d-independent constants $b_0, \underline{\sigma}, \overline{\sigma} > 0$, $j_0 \in \{1,\dots,d\}$, $k_0 \in \{2,\dots,d\}$ and $\alpha' \in (0,1)$, such that

a) $P_{ij} = 0$ for all $1 \le i,j \le d$ such that $|j-i| > j_0$.

b) $\sum_{i=1}^d P_{ij} \le \alpha'$ for all $1 \le j \le d$.

c) $\underline{b}^{(k)} \ge b_0$ for $k_0 \le k \le d$.

d) $\sigma_i \in [\underline{\sigma}, \overline{\sigma}]$ for $1 \le i \le d$.

We check that these conditions imply Assumption 3.2.2 with $r^* = 0$. Recall that the only difference between Assumption 3.2.1 and Assumption 3.2.2 is II in the former and II' in the latter. Note c) and d) immediately imply III, IV of Assumption 3.2.1 with $r^* = 0$ and $b_0, \underline{\sigma}, \overline{\sigma}$ as above.

Condition b) and induction imply

$$\max_{1\le i,j\le d} P_{ij}^n \le \max_{1\le j\le d} \sum_{i=1}^{d} P_{ij}^n \le \left(\max_{1\le i,l\le d} P_{il}^{n-1}\right) \max_{1\le j\le d} \sum_{l=1}^{d} P_{lj} \le (\alpha')^n, \qquad n \ge 1. \tag{3.2.16}$$

Therefore, since $R^{-1} = \sum_0^\infty \left(P^T\right)^n$, condition II' holds with $M = 1/(1-\alpha')$. It remains only to show I of Assumption 3.2.1. To simplify the proof we suppose $j_0 = 1$; the general case is similar. Consider i, j such that $j > i$. Then, by part (a) of the above assumptions, $P_{ji}^n = 0$ for $n < j - i$. This fact and (3.2.16) give

$$R_{ij}^{-1} = \sum_{n=0}^{\infty} (P^T)_{ij}^n = \sum_{n=j-i}^{\infty} P_{ji}^n \le \sum_{n=j-i}^{\infty} (\alpha')^n = \frac{(\alpha')^{j-i}}{1-\alpha'}. \tag{3.2.17}$$

This proves I of Assumption 3.2.1 with $\alpha = \alpha'$ and $C = 1/(1-\alpha')$. The case where $j_0 > 1$ is proven similarly, with $\alpha = (\alpha')^{1/j_0}$ and C being a dimension-independent multiple of $1/(1-\alpha')$. Applying these facts to Theorem 3.2.1 in the case of Assumption 3.2.2 with $r^* = 0$ gives the following theorem.

Theorem 3.2.3. *Suppose X satisfies a) to d) of Example 3.2.2 and recall $\mathcal{S}(b,\cdot)$ from* (3.2.8). *Then there exist constants $\bar{C}, \bar{C}_0, t_0' > 0$ not depending on d such that for any $B \in (0,\infty)$, $x \in \mathcal{S}(b, B)$ and $d > t_0'$,*

$$\begin{aligned} &\mathbb{E}\left[\|X(x,t) - X(X(\infty),t)\|_{1,(\alpha')^{1/2j_0}}\right] \\ &\qquad\qquad \le \bar{C}\left(\sqrt{1+t} + \|x\|_\infty e^{B/\underline{\sigma}^2}\right) e^{-\bar{C}_0 \frac{t}{\log(t\wedge d)}}, \qquad t \ge t_0'. \end{aligned} \tag{3.2.18}$$

3.3 Perturbations from stationarity for the Symmetric Atlas Model

This section is dedicated to the study of dimension free convergence for the Symmetric Atlas model, namely the model defined in (3.2.11) with $p = 1/2$. We view this model as a first step to explore cases in which Assumption 3.2.1 fails to hold. As opposed to stretched exponential convergence rates obtained in Section 3.2, we obtain dimension-free convergence rates to stationarity for the process at a *polynomial rate* if started from appropriate perturbations from stationarity.

Recall that the gap process X of the Symmetric Atlas model has the law of $\mathrm{RBM}(\Sigma, \mu, R)$ where $\mu = -(1, 0 \ldots, 0)$, $R = I - P^T$ and $\Sigma = 2R$ for

$$P_{ij} = \begin{cases} 1/2 & j = i+1, \\ 1/2 & j = i-1, \\ 0 & \text{otherwise}. \end{cases} \tag{3.3.1}$$

R^{-1} is given by computation (e.g. Banerjee and Budhiraja (2020, Proof of Theorem 4), or by taking $p \to 1/2$ in (3.2.13):

$$R^{-1}_{ij} = \begin{cases} 2i\left(1 - \frac{j}{d+1}\right) & 1 \leq i \leq j \leq d, \\ 2j\left(1 - \frac{i}{d+1}\right) & 1 \leq j < i \leq d. \end{cases} \tag{3.3.2}$$

The above representation shows that R^{-1} violates I, II of Assumption 3.2.1, for example by considering $i = j = \lfloor d/2 \rfloor$. Nonetheless, $b = -R^{-1}\mu = \{R^{-1}_{i1}\}_{i=1}^d > 0$ and $\Sigma_{ii} = 2$ for all i. Therefore, there exists a stationary distribution. In fact, if $X(\infty)$ denotes the corresponding stationary distributed random variable, it holds that Harrison and Williams (1987b); Ichiba et al. (2011)

$$X(\infty) \sim \bigotimes_{i=1}^{d} \mathrm{Exp}\left(2\left(1 - \frac{i}{d+1}\right)\right). \tag{3.3.3}$$

3.3.1 Main result and applications

Though Theorem 3.2.1 does not hold, we employ different methods to obtain dimension-free convergence rates to stationarity from initial conditions that perturb the stationary distribution by random variables in a 'perturbation class', which we now define.

Definition 3.3.1 (Perturbation Class). *For $P_1, P_2, \delta \in (0,\infty)$, let $\mathcal{P}(P_1, P_2, \delta)$ denote the class of $\mathbb{R}^\infty$-valued random vectors $Y = (Y_1, Y_2, \dots)$ satisfying the following:*

(i) $\mathbb{E}\left[\|Y\|_1^2\right] \le P_1$.

(ii) $\sup_{m\in\mathbb{N}} \mathbb{E}\left[\exp\left\{\delta m^{-2} \|Y|_m\|_\infty\right\}\right] \le P_2$.

We will consider synchronously coupled processes, one starting from stationarity and the other starting from a perturbation of this stationary configuration by a random vector in $\mathcal{P}(P_1, P_2, \delta)$ for some $P_1, P_2, \delta \in (0,\infty)$. Define for $Y \in \mathcal{P}(P_1, P_2, \delta)$

$$\alpha^Y(n) := \mathbb{E}\left[\sum_{i=n+1}^{\infty} |Y_i|\right], \quad n \in \mathbb{N}. \tag{3.3.4}$$

By assumption (i) above on the class $\mathcal{P}(P_1, P_2, \delta)$, note that for any $Y \in \mathcal{P}(P_1, P_2, \delta)$, $\alpha^Y(n) \to 0$ as $n \to \infty$.

Theorem 3.3.2. *Fix any $P_1, P_2, \delta \in (0,\infty)$ and $Y \in \mathcal{P}(P_1, P_2, \delta)$. Let $X(\infty)$ be distributed as in* (3.3.3) *and define $X^Y(\infty) := (X(\infty) + Y|_d)_+$.*

Then, there exist constants $t_0, t_0'', C_0, C_1 \in (0,\infty)$ not depending on P_1, P_2, δ such that for any $d \ge 1$ and any $n : \mathbb{R}_+ \to \mathbb{N}$ satisfying $\alpha^Y(n(t)) \to 0$ and $t^{-3/32} n(t) \to 0$ as $t \to \infty$,

$$\begin{aligned}
&\mathbb{E}\left[\|X(X^Y(\infty), t) - X(X(\infty), t)\|_1\right] \\
&\quad \le \begin{cases} C_1\sqrt{P_1}\, n(t) t^{-3/32} + C_1\sqrt{P_1}\left(1 + P_2^{1/4}\right) e^{-C_0 \frac{\delta}{\delta+4} t^{3/16}} + \alpha^Y(n(t)), & t \in \left[t_0^{(n)},\, d^{16/3}\right), \\ C_1\sqrt{P_2(d^2 + P_1)}\, e^{-C_0 \frac{t}{d^6 \log(2d)}}, & t \in \left[t_0''\, d^4 \log(2d),\, \infty\right), \end{cases}
\end{aligned} \tag{3.3.5}$$

where $t_0^{(n)} := \inf\{t \ge t_0 : t^{3/16} \ge 1 + 2n(t)\}$.

Remark 3.3.1. *Note that the bounds in Theorem 3.3.2 show polynomial decay when $t < d^{16/3}$ and exponential decay for $t > d^6 \log(2d)$. In particular, we do not obtain the 'smooth patching' of the bounds as in the results of Section 3.2. This is mainly because the methods used for the two regimes $t < d^{16/3}$ and $t > d^6 \log(2d)$ in Theorem 3.3.2 are starkly different. The 'contractions' in $\|\cdot\|_{1,\beta}$ distance between the coupled RBMs upon certain events taking place in their trajectory,* *which was key to the results in Section 3.2, no longer holds here due to Assumption 3.2.1 not being satisfied. This is the main factor behind the discontinuous qualitative and quantitative transitions between the bounds in the two regimes in Theorem 3.3.2. See also Remark 3.3.2.*

The choice of $n(\cdot)$ in Theorem 3.3.2 has been intentionally kept flexible. One can choose $n(\cdot)$ in an 'optimal' way so as to minimize $\max\{n(t)t^{-3/32}, \alpha^Y(n(t))\}$. This, in turn, is intricately tied to the distributional behavior of the perturbation vector Y as quantified by the function $\alpha^Y(\cdot)$. We mention the following two special cases as corollaries and choose $n(\cdot)$ in a case-specific way.

For perturbations from stationarity by finitely many coordinates in the following sense, one can take $n(\cdot)$ to be the (fixed) number of perturbed coordinates to obtain the following simplified bound.

Corollary 3.3.3 (Finite perturbations from stationarity)**.** *Fix an integer $m \geq 1$ and a random vector $Z \in \mathbb{R}^m$ such that its extension to $\mathbb{R}^\infty$ given by $Y = (Z, 0, \ldots)$ is in the class $\mathcal{P}(P_1, P_2, \delta)$ of Definition 3.3.1 for some $P_1, P_2, \delta \in (0, \infty)$. Setting $n(t) = m$ for all t, we have for all $d > 1 + 2m$,*

$$
\begin{aligned}
&\mathbb{E}\left[\|X(X^Y(\infty), t) - X(X(\infty), t)\|_1\right] \\
&\quad \leq \begin{cases} C_1\sqrt{P_1}mt^{-3/32} + C_1\sqrt{P_1}\left(1 + P_2^{1/4}\right)e^{-C_0\frac{\delta}{\delta+4}t^{3/16}}, & t_0 \vee (1+2m)^{16/3} \leq t < d^{16/3}, \\ C_1\sqrt{P_2(d^2 + P_1)}e^{-C_0\frac{t}{d^6\log(2d)}}, & t \geq t_0''\, d^4\log(2d). \end{cases}
\end{aligned}
$$

The following corollary addresses the special case of perturbations from stationarity by independent exponential random variables.

Corollary 3.3.4 (Independent exponential perturbations)**.** *Consider $Y = (Y_1, Y_2, \ldots)$ where $\{Y_i\}_{i\geq 1}$ are independent random variables with $Y_i \sim Exp(i^{1+\beta})$ (exponential with mean $i^{-(1+\beta)}$), for some $\beta > 0$. Then $Y \in \mathcal{P}(P_1, P_2, \delta)$ with $P_1 := \sum_1^\infty i^{-2(1+\beta)} + \left(\sum_1^\infty i^{-(1+\beta)}\right)^2$, $P_2 :=$*

$1+\sum_1^\infty i^{-(1+\beta)}$ *and* $\delta := 1/2$*. Setting* $n(t) = \lfloor t^{\frac{3}{32(1+\beta)}} \rfloor$*, we have*

$$\mathbb{E}\left[\|X(X^Y(\infty),t) - X(X(\infty),t)\|_1\right]$$
$$\leq \begin{cases} \left(C_1\sqrt{P_1} + \frac{2}{\beta}\right) t^{-\frac{\beta}{1+\beta}\frac{3}{32}} + C_1\sqrt{P_1}\left(1+P_2^{1/4}\right)\exp\left\{-\frac{C_0}{9}t^{3/16}\right\}, & t_0' \leq t < d^{16/3}, \\ \\ C_1\sqrt{P_2(d^2+P_1)}\exp\left(-C_0\frac{t}{d^6\log(2d)}\right), & t \geq t_0''\ d^4\log(2d), \end{cases}$$

where $t_0' \in (0,\infty)$ *does not depend on* d *or* β*.*

The proof of this corollary makes clear one could consider independent $Y_i \sim Exp(\lambda_i)$ for any sequence $\{\lambda_i\}_{i\geq 1}$ such that $\|Y\|_1$ has finite expectation and variance. We choose $\lambda_i = i^{1+\beta}$ as it lends itself to simple and explicit calculations of the rates of convergence.

Proof of Corollary 3.3.4. $Y \in \mathcal{P}(P_1, P_2, \delta)$ is the result of the following calculations:

$$\mathbb{E}\left[\|Y\|_1^2\right] = \mathrm{Var}(\|Y\|_1) + \left(\mathbb{E}\left[\|Y\|_1\right]\right)^2 = \sum_{i=1}^\infty i^{-2(1+\beta)} + \left(\sum_{i=1}^\infty i^{-(1+\beta)}\right)^2,$$
$$\mathbb{E}\left[\exp\left\{\frac{\|Y|_m\|_\infty}{2m^2}\right\}\right] \leq 1 + m^{-2}\sum_{i=1}^m i^{-(1+\beta)} \leq 1 + \sum_{i=1}^\infty i^{-(1+\beta)}, \quad \text{for all } m \in \mathbb{N}.$$

With $n(t) = \lfloor t^{\frac{3}{32(1+\beta)}} \rfloor$ we have by basic calculus that $\alpha^Y(n(t)) \leq \frac{2}{\beta} t^{-\frac{\beta}{1+\beta}\frac{3}{32}}$ and $n(t)t^{-\frac{3}{32}} \leq t^{-\frac{\beta}{1+\beta}\frac{3}{32}}$ for $t \geq 2$. Applying Theorem 3.3.2 gives the corollary. □

Remark 3.3.2. *In Theorem 3.3.2 and Corollaries 3.3.3 and 3.3.4, the upper bound has a polynomial decay in* t *for large* d *(for* $t < d^{16/3}$*) as opposed to the stretched exponential decay observed in Section 3.2 when Assumption 3.2.1 applies. Although we do not currently have associated lower bounds, we strongly believe that the* L^1*-Wasserstein distance of the perturbed system (as defined in Theorem 3.3.2) from stationarity indeed shows polynomial decay for the Symmetric Atlas model. This belief stems from the dynamics of the associated killed Markov chain whose transition kernel is prescribed by* P *(see discussion after Assumption 3.2.1) which are shown throughout this chapter to govern convergence rates to stationarity. This Markov chain for the Symmetric Atlas model behaves as a simple random walk away from the cemetery* *state and thus lacks the 'strong drift' towards the cemetery state characteristic of the models considered in Section 3.2. This results in the slower convergence rates.*

The polynomial rates of convergence to stationarity obtained in Banerjee and Burdzy (2020) for the Potlatch process on $\mathbb{Z}^k$, which (for $k = 1$) can be loosely thought of as a 'Poissonian version' of the gap process of the infinite Symmetric Atlas model constructed in Pal and Pitman (2008), lends further evidence to this belief.

3.3.2 A pathwise derivative approach towards convergence rates

The proof of Theorem 3.3.2 is based on an analysis of the derivative process (derivative taken with respect to initial conditions) of the RBM X. The key observation made here is a representation of this derivative process in terms of a *random walk in a certain random environment* constructed from the random order in which the RBM hits distinct faces of the orthant $\mathbb{R}_+^d$ (see (3.3.8)). This representation, in turn, is based on a succinct form for the derivative process obtained in Andres (2009, Theorem 1.2). This is summarized in Theorem 3.3.6 below. This representation is interesting in its own right and we believe a systematic study of the derivative process is at the heart of obtaining convergence rates in more general cases where Assumption 3.2.1 does not hold. Moreover, as the relationship between the derivative process and the (random) transition kernel of the random walk in the random environment is an exact equality (3.3.8), this representation should also lead to *lower bounds for convergence rates.* We hope to report on this in future work.

In the probability literature, random walks in random environments most commonly appear as random walks on graphs with jump probabilities given by i.i.d. random variables (see e.g. Sznitman, A.-S. (2004) or Dembo et al. (2004) for a model with i.i.d. holding times). Since the process we will consider in Theorem 3.3.6 is substantially different, we take some care first to define it.

Definition 3.3.5 ($RW(\mathbf{a}, i_0)$)**.** *Here we define a random walk on $\{0, \ldots, d+1\}$, for $d \geq 1$, in a given fixed environment* $\mathbf{a}$ *and initial condition i_0. Call any sequence* $\mathbf{a} := (l_k, t_k)_{k \geq 0}$ admissible *if*

(i.) $(l_k, t_k) \in \{1, \ldots, d\} \times [0, \infty)$ *for all* $k \geq 0$,

(ii.) $t_0 = 0$ *and* $\{t_k\}_{k\geq 0}$ *is strictly increasing.*

For any admissible sequence $\mathbf{a}$ *and any* $i \in \{1, \ldots, d\}$, *define the projected admissible sequence* $\mathbf{a}_i = (l^i_k, t^i_k)_{k\geq 0} = (i, t^i_k)_{k\geq 0}$ *to be the unique admissible sequence obtained from the elements of the set* $\mathbf{a} \cap \{\{i\} \times [0, \infty)\}$.

Define the random walk in environment $\mathbf{a}$ *started from* $i_0 \in \{0, \ldots, d+1\}$, *written as* $RW(\mathbf{a}, i_0)$, *to be the time-inhomogeneous Markov process* W *with state space* $\{0, \ldots, d+1\}$ *whose law is uniquely characterized by the following:*

(i.) $W(0) = i_0$,

(ii.) W *is absorbed at* 0 *and* $d+1$,

(iii.) *With* $\{T_k\}_{k\geq 0} = \{T_k(\mathbf{a}, i_0)\}_{k\geq 1}$ *defined by* $T_0 = 0, T_1 = t^{i_0}_1$ *and*

$$T_{k+1} = \min\left\{t^i_j \,:\, i = W(T_k),\ t^i_j > T_k,\ (i, t^i_j) \in \mathbf{a}\right\}, \qquad k \geq 1,$$

we have

$$\begin{aligned}&\mathbb{P}_{\mathbf{a},i_0}\left(W(T_{k+1}) = W(T_k) + 1 \mid (W(T_k), T_k)\right) \\ &\qquad\qquad = \mathbb{P}_{\mathbf{a},i_0}\left(W(T_{k+1}) = W(T_k) - 1 \mid (W(T_k), T_k)\right) = 1/2.\end{aligned}$$

(iv.) $\mathbb{P}_{\mathbf{a},i_0}\left(W(t) = W(T_k) \mid (W(T_k), T_k)\right) = 1$ *for* $t \in [T_k, T_{k+1})$, $k \geq 0$,

(v.) *for* $0 \leq t < t'$,

$$\mathbb{P}_{\mathbf{a},i_0}\big(W(t') = W(t) \mid W(t) = 0\big) = \mathbb{P}_{\mathbf{a},i_0}\big(W(t') = W(t) \mid W(t) = d+1\big) = 1.$$

In the above, we used the suffix in the probabilities to highlight the dependence of the law of W *on* $\mathbf{a}$ *and* i_0. *The process* W *can be seen as a simple random walk absorbed at* $0, d+1$ *with jump times prescribed by the points in* $\mathbf{a}$ *encountered along its trajectory.*

Finally, define

$$J_{\mathbf{a},i_0}(t) := \#\left\{s \in [0,t] \,:\, W(s-) \neq W(s)\right\} = \#\left\{k \geq 1 : T_k \in [0,t]\right\},$$

to be the number of jumps made by $RW(\mathbf{a}, i_0)$ in the time interval $[0,t]$.

We now define a few additional conventions and notations required to state the theorem. For two vectors $x, y \in \mathbb{R}^d$ we write $\langle x, y\rangle$ for the standard inner product, and $e_i, 1 \le i \le d$ for the standard basis vectors. For a $d \times d$ matrix R, write $R^{(i)}$ for the i-th column vector of R.

For X started at $x \in \mathbb{R}^d_+$, $x > 0$, define a sequence of stopping times as follows: $\tau_0(x) = 0, \tau_1(x) = \inf\{t > 0 \mid X_i(x,t) = 0 \text{ for some } i\}$ and for $k \ge 1$,

$$\tau_{k+1}(x) = \inf\{t > \tau_k(x) \mid X_i(x,t) = 0,\ X_j(x,\tau_k) = 0 \text{ for some } i,j \text{ such that } j \neq i\}. \quad (3.3.6)$$

Also define the sequence of integers $i_k(x)$ for $k \ge 0$ as follows: Fix any $i_0(x) \in \{1,\dots,d\}$ and define the remaining $i_k(x)$ by $X_{i_k}(x,\tau_k) = 0$, i.e. $i_k(x)$ is the index of the coordinate hitting zero at time $\tau_k(x)$ for $k \ge 1$. In other words, $\{\tau_k(x)\}_{k\ge1}$ represent the times when X has crossed from one face of the orthant to another, and i_k tells which coordinate has hit zero at crossing time τ_k. We suppress dependence of τ_k, i_k on x when there is no risk of confusion.

From Sarantsev (2015, Theorem 1.9), the Atlas model almost surely has no triple or simultaneous collisions and thus, almost surely, for any $x \in \mathbb{R}^d_+, t > 0$, $X_i(x,t) = 0$ for at most one $i \in \{1,\dots,d\}$. Therefore, i_k, τ_k are well-defined and the sequence $\{(i_k,\tau_k)\}_{k\ge0}$ is admissible in the sense of Definition 3.3.5. This fact is essential for the random walk representation below.

Theorem 3.3.6. *For every $t \in [0,\infty)$ and every $x > 0$, the map $y \mapsto X(y,t)$ is almost surely differentiable at x. For each $i_0 \in \{1,\dots,d\}$ the process $\eta^{i_0}(x,t) := \lim_{\epsilon\to0}\epsilon^{-1}(X(x+\epsilon e_{i_0},t) - X(x,t))$ has a right-continuous modification defined on $[0,\infty)$ such that*

$$\eta^{i_0}(x,t) = S^{i_0}_k(x), \qquad \text{for } t \in [\tau_k,\tau_{k+1}), \quad k \ge 0, \quad (3.3.7)$$

where $\left\{S^{i_0}_k(x)\right\}_{k\ge0}$ is a sequence of d-dimensional random vectors iteratively defined by

$$\begin{cases} S^{i_0}_0(x) = e_{i_0} \\ S^{i_0}_{k+1}(x) = S^{i_0}_k(x) - \langle S^{i_0}_k(x), e_{i_{k+1}}\rangle R^{(i_{k+1})}, \quad k \ge 0. \end{cases}$$

Moreover, $\Theta(x) := \{(\tau_k, i_k)\}_{k\geq 0}$ *is admissible and the derivative process has the following representation in terms of the law of* $RW(\Theta(x), i_0)$*:*

$$\eta_j^{i_0}(t,x) = \mathbb{P}_{\Theta(x),i_0}(W(t) = j), \; j = 1, \dots, d. \tag{3.3.8}$$

Remark 3.3.3. *We clarify the relationship between boundary-hitting times of the process* X *started at* $x > 0$ *and the jump times of* $W \sim RW(\theta(x), i_0)$*,* $i_0 \in \{1, \dots, d\}$*.*

Suppose X *begins at* $x > 0$ *and* $W(t) = i \in \{1, \dots, d\}$ *at some time* $t \geq 0$*. Then at the first time after* t *that* X_i *hits zero,* W *will jump to* $i - 1$ *or* $i + 1$ *with equal probability.*

Now suppose for a given time interval $[0, T]$ *and integer* $m \geq 2$ *the random walk* W *starting from* i_0 *remains in the set* $\{1, \dots, m-1\}$*. Suppose also that in* $[0, T]$ *each of the first* m *coordinates of the process* X *have hit zero at least* $N \geq 1$ *times. Then the walk has made at least* N *jumps in the time interval* $[0, T]$*. Thus, recalling* $\mathcal{N}_m(x,t)$ *from* (3.4.2)*,*

$$\begin{aligned} &\{\mathcal{N}_m(x,t) \geq N, \; W(s) \in \{1, \dots, m-1\} \quad \textit{for} \quad s \in [0,t]\} \\ &\qquad \subseteq \left\{J_{\Theta(x),i_0}(t) \geq N, \; W(s) \in \{1 \dots m-1\} \quad \textit{for} \quad s \in [0,t]\right\}. \end{aligned} \tag{3.3.9}$$

This fact will be crucially used in the proof of Theorem 3.3.2.

We also note here that the process W *is non-standard in the sense that the number of jumps of* W *in a certain time interval depends on the whole trajectory of* W *in that interval, which makes its analysis challenging.*

Remark 3.3.4. *We have stated Theorem 3.3.6 for the Symmetric Atlas model examined here, but an analogous result holds for any RBM* (3.1.1) *that almost surely does not hit intersections of faces (corners) of the orthant* $\mathbb{R}_+^d$*. In that case one-step transitions are given by the matrix* P *(from* $R = I - P^T$*). See Karatzas, I. and Pal, S. and Shkolnikov, M. (2016) for conditions guaranteeing when the gap process of an Atlas model (symmetric or asymmetric) does not hit corners, and Sarantsev (2015) for similar conditions for a general RBM.*

For the general RBM (3.1.1)*, even when corners are hit with positive probability, Mandelbaum, A. and Ramanan, K. (2010) shows that the derivative process exists in an appropriate sense. However, in the general case we do not have a random walk representation as in Theorem*

3.3.6. Blanchet et al. (2020) has recently obtained an upper bound for the derivative process in terms of products of random matrices derived in terms of the boundary hitting times and locations of the RBM and the killed Markov process associated with P (see Blanchet et al. (2020, Lemma 5)). This presents an opportunity to generalize the methods used here, and we defer it to future work.

The following corollary to Theorem 3.3.6 is the key tool in proving Theorem 3.3.2.

Corollary 3.3.7. *Fix $x, \tilde{x} \in \mathbb{R}^d_+$ with $x > 0$ and let $\gamma(u) = x + u(\tilde{x} - x)$ for $u \in [0,1]$. Then, writing $\tau^*_0 := \inf\{s \geq 0 : W(s) = 0\}$,*

$$\|X(\tilde{x}, t) - X(x,t)\|_1 \leq \sum_{i=1}^{d} |(\tilde{x} - x)_i| \int_{[0,1)} \mathbb{P}_{\Theta(\gamma(u)),i}(\tau^*_0 > t)\, du, \quad t \geq 0. \tag{3.3.10}$$

Proof. For each $i = 1, \ldots, d$ and $t \geq 0$ define the function $f_{i,t} : [0,1] \mapsto [0, \infty)$ as $f_{i,t}(u) = X(\gamma(u), t)$. As shown in the proof of Harrison and Reiman (1981, Theorem 1), $x \mapsto X_i(x,t)$ is Lipschitz. Thus $f_{i,t}$ is absolutely continuous on $[0,1]$ and we have for $t \geq 0$:

$$\begin{aligned}
\|X(\tilde{x}, t) - X(x,t)\|_1 &\leq \sum_{j=1}^{d} \sum_{i=1}^{d} |(\tilde{x} - x)_i| \int_{[0,1)} \mathbb{P}_{\Theta(\gamma(u)),i}(W(t) = j)\, du \\
&= \sum_{i=1}^{d} |(\tilde{x} - x)_i| \int_{[0,1)} \mathbb{P}_{\Theta(\gamma(u)),i}(W(t) \in \{1, \ldots, d\})\, du \\
&\leq \sum_{i=1}^{d} |(\tilde{x} - x)_i| \int_{[0,1)} \mathbb{P}_{\Theta(\gamma(u)),i}(\tau^*_0 > t)\, du.
\end{aligned}$$

The first step above follows from absolute continuity and Theorem 3.3.6 for $\gamma(u) > 0$ for $u \in [0,1)$. The second step follows by an interchange of summation. □

3.4 Proofs: Dimension-free local convergence rates for RBM

3.4.1 Boundary-hitting times

Before proceeding to the proofs, we define boundary hitting times for a solution X to (3.1.1), which we use throughout. For any $1 \leq d' \leq d$, we define a sequence of times at which X hits d' faces of $\mathbb{R}^d_+$ corresponding to $X_i = 0$ for $i = 1, \ldots, d'$. Set $\eta^0_{d'}(x) = 0$ and define inductively for

$k \geq 1$

$$\xi_i^k(x) = \inf\{t > \eta_{d'}^{k-1}(x) + 1 \mid X_i(x,t) = 0\}, \qquad \eta_{d'}^k(x) = \max\{\xi_i^k(x) \mid i = 1 \dots d'\} \tag{3.4.1}$$

where we suppress the d' dependence of ξ_i^ks for convenience. Also define

$$\mathcal{N}_{d'}(x,t) = \max\{k \mid \eta_{d'}^k(x) \leq t\}. \tag{3.4.2}$$

All the stopping times defined above are finite almost surely, which follows from the positive recurrence criterion $R^{-1}\mu < 0$. It can also be deduced from Lemma 3.4.5 below.

3.4.2 Fundamental properties of RBM

The next two theorems record fundamental results related to this work from, respectively, Kella, O. and Ramasubramanian, S. (2012) Theorem 1.1, and Sarantsev, A. (2019) Theorem 3.1, Corollaries 3.5 and 3.6.

Theorem 3.4.1 (Monotonicity under synchronous coupling)**.** *For X a solution to* (3.1.1) *and $x, \tilde{x} \in \mathbb{R}_+^d$ such that $x \geq \tilde{x}$, the following hold:*

(i) $X(x,t) \geq X(\tilde{x},t)$ for all $t > 0$.

(ii) $t \mapsto L(x,t) - L(\tilde{x},t)$ is non-positive, non-increasing and bounded below by $-R^{-1}(x - \tilde{x})$.

(iii) $t \mapsto R^{-1}(X(x,t) - X(\tilde{x},t)) = R^{-1}(x - \tilde{x}) + L(x,t) - L(\tilde{x},t)$ is non-negative and non-increasing.

Theorem 3.4.2 (Stochastic domination of projected system)**.** *Suppose X is a solution to* (3.1.1) *with parameters (Σ, μ, R) and corresponding local times L. For $x \in \mathbb{R}_+^d$ and an integer $1 \leq k \leq d$, define the process $Z(x|_k, t) := x|_k + \mu|_k t + (DB(t))|_k,\ t \geq 0$, which uses the same driving Brownian motion B as X. Define $\bar{X}$ to be the $\mathbb{R}_+^k$-valued process obtained as the solution to*

$$\bar{X}(x|_k, t) = Z(x|_k, t) + R|_k \bar{L}(x|_k, t),\ t \geq 0,$$

where $\bar{L}(x|_k, \cdot)$ is the local time which constrains $\bar{X}$ to $\mathbb{R}^k_+$. Then

$$X|_k(x,t) \le \bar{X}(x|_k,t) \quad t \ge 0, \qquad L|_k(x,t) - L|_k(x,s) \ge \bar{L}(x|_k,t) - \bar{L}(x|_k,s) \quad 0 \le s \le t.$$

3.4.3 Proofs

The following lemma provides a crucial local contraction estimate. It shows that for any $x \in \mathbb{R}^d_+$, the weighted distance between the coupled processes $X(x,\cdot)$ and $X(0,\cdot)$ as measured by $u(x,\cdot)$ in (3.2.1) decreases by a constant factor if a subset of coordinates of $X(x,\cdot)$ (whose cardinality is determined by the initial distance) hit zero.

Lemma 3.4.3 (Local contraction)**.** *Suppose I, II of Assumption 3.2.1 hold for X, an* RBM(Σ,μ,R)*. Fix an initial condition $X(x,0) = x \ge 0$. With α as in Assumption 3.2.1, fix $\beta \in (\alpha,1)$ and $\delta \in (\beta,1)$. Recall the weighted supremum norm $\|x\|_{\infty,\delta} = \max_{1\le i\le d} \delta^i x_i$, and $u(x,\cdot)$ from* (3.2.1).

Fix $d' \in \{1,\ldots,d\}$. Recall the definition of $\eta^1_{d'} = \eta^1_{d'}(x)$ from (3.4.1).

There exist $C' > 0$ and $\lambda \in (1/2,1)$ not dependent on d, d' or x such that,

(i) if $1 \le d' \le d-1$,

$$u(x,0) \ge C'\|x\|_{\infty,\delta}\,(\beta/\delta)^{d'+1} \quad \Longrightarrow \quad u(x,\eta^1_{d'}) \le \lambda u(x,0). \tag{3.4.3}$$

(ii) if $d' = d$,

$$u(x,\eta^1_d) \le \lambda u(x,0). \tag{3.4.4}$$

C', λ may be chosen explicitly as functions of β, δ and the constants α, C, M from Assumption 3.2.1.

Proof. Define the processes

$$\begin{aligned} \Delta X(t) &= X(x,t) - X(0,t) \\ \Delta L(t) &= L(x,t) - L(0,t) \\ Y(t) &= R^{-1}\Delta X(t) = R^{-1}x + \Delta L(t) \end{aligned} \tag{3.4.5}$$

From Theorem 3.4.1 we know that for all for all $t \geq 0$, $\Delta X(t) \geq 0$, $t \mapsto \Delta L(t)$ is non-positive, non-increasing and $t \mapsto Y(t)$ is non-negative, non-increasing. By definition, then, $t \mapsto u(x,t)$ is non-negative and non-increasing. We aim to show that u indeed contracts by a fixed proportion λ of its initial value at time $\eta^1_{d'}$.

The crucial fact is that if $X_i(x, \cdot)$ has hit zero before a time t, then $\Delta L_i(s) \leq -x_i$ for all $s \geq t$. Indeed, setting $t_0 > 0$ to be the first hitting time of $X_i(x, \cdot)$ at 0 and assuming $t_0 < t$,

$$0 = \Delta X_i(t_0) = x_i + (R\Delta L(t_0))_i = x_i + \Delta L_i(t_0) - \left(P^T \Delta L(t_0)\right)_i \geq x_i + \Delta L_i(t) \geq x_i + \Delta L_i(s), \tag{3.4.6}$$

for all $s \geq t$, where the first equality follows from $R = I - P^T$ and the last two inequalities follow from Theorem 3.4.1 (ii) and the non-negativity of P.

By definition, at time $\eta^1_{d'} = \eta^1_{d'}(x)$ the first d' coordinates of $X(x, \cdot)$ have already hit zero. (3.4.6) then implies

$$\begin{aligned} u(x, \eta^1_{d'}) &= \sum_{i=1}^{d} \beta^i Y_i(\eta^1_{d'}) = u(x,0) + \sum_{i=1}^{d} \beta^i \Delta L_i(\eta^1_{d'}) \\ &\leq u(x,0) - \sum_{i=1}^{d'} \beta^i x_i + \mathbb{1}_{d'<d} \sum_{i=d'+1}^{d} \beta^i \Delta L_i(\eta^1_{d'}) \leq u(x,0) - \sum_{i=1}^{d'} \beta^i x_i. \end{aligned} \tag{3.4.7}$$

The last inequality follows once again from Theorem 3.4.1 (ii). To achieve the result (3.4.3), we first bound $\sum_{i=d'+1}^{d} \beta^i Y_i(0)$. In the following, the first inequality is a consequence of the definition of $\|x\|_{\infty,\delta}$ and the second inequality follows from I, II of Assumption 3.2.1. Remaining statements follow from the fact that $\alpha < \beta < \delta < 1$. For $d' < d$,

$$\begin{aligned}\sum_{i=d'+1}^{d} \beta^i Y_i(0) &= \sum_{i=d'+1}^{d} \beta^i \sum_{j=1}^{d} R^{-1}_{ij} x_j \le \|x\|_{\infty,\delta} \sum_{i=d'+1}^{d} \beta^i \sum_{j=1}^{d} R^{-1}_{ij} \delta^{-j} \\ &\le \|x\|_{\infty,\delta} \sum_{i=d'+1}^{d} \beta^i \left(M \sum_{j=1}^{i} \delta^{-j} + C \sum_{j=i+1}^{d} \alpha^{j-i}\delta^{-j} \right) \\ &\le \|x\|_{\infty,\delta} \sum_{i=d'+1}^{d} (\beta/\delta)^i \left(M \sum_{j=0}^{i-1} \delta^{j} + C \sum_{j=i+1}^{\infty} (\alpha/\delta)^{j-i} \right) \\ &\le \frac{\|x\|_{\infty,\delta} M}{1-\delta} \sum_{i=d'+1}^{d} (\beta/\delta)^i + \frac{\|x\|_{\infty,\delta}\, C(\alpha/\delta)}{1-\alpha/\delta} \sum_{i=d'+1}^{d} (\beta/\delta)^i \le \tilde{C}\|x\|_{\infty,\delta}\, (\beta/\delta)^{d'+1},\end{aligned} \tag{3.4.8}$$

with $\tilde{C} = \frac{M}{(1-\delta)(1-\beta/\delta)} + \frac{C(\alpha/\delta)}{(1-\alpha/\delta)(1-\beta/\delta)}$, which by Assumption 3.2.1 does not depend on d', d or x.

Now recall that since P is transient and $R = I - P^T$ we have $R^{-1} = \sum_{n=0}^{\infty}(P^T)^n$, which implies $Y(0) = R^{-1}x \ge x$. Using this and (3.4.8), we have for $1 \le d' \le d-1$

$$\sum_{i=1}^{d'} \beta^i x_i = \sum_{i=1}^{d} \beta^i x_i - \sum_{i=d'+1}^{d} \beta^i x_i \ge \sum_{i=1}^{d} \beta^i x_i - \sum_{i=d'+1}^{d} \beta^i Y_i(0) \ge \sum_{i=1}^{d} \beta^i x_i - \tilde{C}\|x\|_{\infty,\delta}\, (\beta/\delta)^{d'+1}. \tag{3.4.9}$$

Furthermore, I, II of Assumption 3.2.1 and $1 > \beta > \alpha$ give

$$\begin{aligned}u(x,0) = \sum_{i=1}^{d} \beta^i Y_i(0) = \sum_{j=1}^{d} \beta^j x_j \sum_{i=1}^{d} R^{-1}_{ij} \beta^{i-j} &\le \sum_{j=1}^{d} \beta^j x_j \left(C \sum_{i=1}^{j} (\alpha/\beta)^{j-i} + M \sum_{i=j+1}^{d} \beta^{i-j} \right) \\ &\le \left(\frac{C}{1-\alpha/\beta} + \frac{M\beta}{1-\beta} \right) \sum_{i=1}^{d} \beta^j x_j \le \tilde{C}' \sum_{i=1}^{d} \beta^j x_j,\end{aligned} \tag{3.4.10}$$

where we have set $\tilde{C}' = 1 \vee [C/(1-\alpha/\beta) + M\beta/(1-\beta)]$. Combining (3.4.9) and (3.4.10),

$$\sum_{i=1}^{d'} \beta^i x_i \ge \frac{1}{\tilde{C}'} u(x,0) - \tilde{C}\|x\|_{\infty,\delta}\, (\beta/\delta)^{d'+1}. \tag{3.4.11}$$

Finally, if $u(x,0) \geq 2\tilde{C}'\tilde{C}\|x\|_{\infty,\delta}\,(\beta/\delta)^{d'+1}$ then (3.4.11) gives

$$\sum_{i=1}^{d'} \beta^i x_i \geq \frac{1}{2\tilde{C}'} u(x,0). \tag{3.4.12}$$

The result (3.4.3) now follows with $C' = 2\tilde{C}'\tilde{C}$ and $\lambda = 1 - 1/(2\tilde{C}')$ using (3.4.12) and (3.4.7). To prove (3.4.4), we use (3.4.7) with $d' = d$ and (3.4.10) as follows

$$u(x,\eta_d^1) \leq u(x,0) - \sum_{i=1}^{d} \beta^i x_i \leq \left(1 - \frac{1}{\tilde{C}'}\right) u(x,0) \leq \lambda u(x,0). \tag{3.4.13}$$

□

Corollary 3.4.4. *Retain the assumptions of Lemma 3.4.3 and recall β, δ chosen there. Recall the definition of $N_{d'}(x,t)$ from (3.4.2). Define the stopping times with C' as in (3.4.3),*

$$\tau(x,d') := \inf\left\{s > 0 \ \mid\ u(x,s) \leq C'\|x\|_{\infty,\delta}\,(\beta/\delta)^{d'+1}\right\}, \qquad \textit{for } x \in \mathbb{R}_+^d,\ 1 \leq d' \leq d-1. \tag{3.4.14}$$

Then for any $q > 0$,

(i) if $1 \leq d' \leq d-1$,

$$u(x,t)\ \mathbb{1}_{\tau(x,d')>t,\ N_{d'}(x,t)\geq q} \leq \lambda^{\lfloor q \rfloor} u(x,0). \tag{3.4.15}$$

(ii) if $d' = d$,

$$u(x,t)\ \mathbb{1}_{N_d(x,t)\geq q} \leq \lambda^{\lfloor q \rfloor} u(x,0). \tag{3.4.16}$$

Proof. First, by Theorem 3.4.1 (iii) and the definition (3.2.1) of $u(x,t)$ we have $u(x,t) \leq u(x,0)$ for all $t > 0$. Therefore, it suffices to show for each $k \geq 1$

$$\begin{aligned} u(x,\eta_{d'}^k) \geq C'(\beta/\delta)^{d'+1} \quad &\Longrightarrow \quad u(x,\eta_{d'}^{k+1}) \leq \lambda u(x,\eta_{d'}^k), \qquad \text{if } 1 \leq d' \leq d-1, \\ \text{and} \quad &\quad\ u(x,\eta_d^{k+1}) \leq \lambda u(x,\eta_d^k). \end{aligned} \tag{3.4.17}$$

To do so, we note that the argument proving Lemma 3.4.3 remains valid if we replace $u(x,\eta_{d'}^1)$ with $u(x,\eta_{d'}^{k+1})$, $u(x,0)$ with $u(x,\eta_{d'}^k)$ and $\Delta X(0) = x$ with $\Delta X(\eta_{d'}^k)$ throughout—so long as (3.4.8) is replaced by $\sum_{i=d'+1}^{d} \beta^i Y_i(\eta_{d'}^k) \leq \sum_{i=d'+1}^{d} \beta^i Y_i(0) \leq \tilde{C}\|x\|_{\infty,\delta}\,(\beta/\delta)^{d'+1}$ in the case where $1 \le d' \le d-1$. This follows directly from (3.4.8) and Theorem 3.4.1 (iii) which gives $Y_i(\eta^k_{d'}) \le Y_i(0)$ for $i = 1, \dots, d$. $\square$

In the following lemma, we obtain estimates on tail probabilities for $\mathcal{N}_{d'}(x,t)$, defined in (3.4.2), using results from Banerjee and Budhiraja (2020) and the stochastic domination recorded in Theorem 3.4.2. Recall k_0 from Assumption 3.2.1, which by definition was such that $d \ge k_0$.

Lemma 3.4.5 (Boundary-hitting estimates)**.** *Fix $d' \in \{k_0, \dots, d\}$. Suppose $\underline{b}^{(d')} > 0$ and IV of Assumption 3.2.1 holds, and recall the definition of $a^{(d')}$ from* (3.2.5)*. Define the d'-dependent quantities*

$$T^{(d')} = 1 + \left(a^{(d')}\right)^2 \log(2d'), \quad \Lambda^{(d')} = \left(a^{(d')}\right)^{-2}. \tag{3.4.18}$$

There exist positive constants δ', C'' and $A_0 \ge 1$ not dependent on d', d, μ, R, Σ, such that for any $x \in \mathbb{R}^d_+$, $A \ge A_0$ and $t \ge 4T^{(d')}/\delta'$,

$$\begin{aligned}\mathbb{P}\left[\mathcal{N}_{d'}(x,t) < \delta' t/(4T^{(d')})\right] &\le \exp\left(-t\frac{\delta' C''}{T^{(d')}}\right) \\ &\quad + \exp\left(-t\frac{C''\Lambda^{(d')}}{A}\right)\left\{1 + \exp\left(\frac{\|x|_{d'}\|_\infty}{A\underline{\sigma}a^{(d')}}\right)\right\}. \end{aligned}\tag{3.4.19}$$

Proof. Define $\bar{X}$ as in Theorem 3.4.2 with $k = d'$. The theorem states $\bar{X}$ dominates $X|_{d'}$, the projection of the d-dimensional RBM with parameters (Σ, μ, R) onto the first d' coordinates. Therefore, a coordinate of $X|_{d'}$ hits zero whenever the same coordinate of $\bar{X}$ hits zero. In other words, $\mathcal{N}_{d'}(x,t)$ dominates the corresponding quantity for $\bar{X}$, for all x, t.

By hypothesis of the lemma, $b^{(d')} > 0$. As in Banerjee and Budhiraja (2020), for any $v \in \mathbb{R}^{d'}_+$ satisfying $R^{-1}v \le b^{(d')}, v > 0$, and any $y \in \mathbb{R}^{d'}_+$, define

$$\|y\|^\star_{\infty,v} := \sup_{1\le i\le d'} v_i\sigma_i^{-2}y_i, \; \Lambda(v) := \inf_{1\le i\le d'} \sigma_i^{-2}v_i^2, \; T(v) := \left(1 + \frac{\log\left(2\sum_{i=1}^{d'} v_i^2\sigma_i^{-2}/\Lambda(v)\right)}{\Lambda(v)}\right).$$

With these definitions, recalling the stochastic domination noted in the previous paragraph, Banerjee and Budhiraja (2020, Proof of Lemma 8, Equations (33) and (41)) applied to the process $\bar{X}$ give positive constants δ', A_0, not depending on d', d, μ, R, Σ, such that for each $x \in \mathbb{R}_+^d$, $A \geq A_0$ and $t \geq 4T(v)/\delta'$,

$$\mathbb{P}\left[\mathcal{N}_{d'}(x,t) < \delta' t/(4T(v))\right] \leq \exp\left(-\frac{\delta' t}{128T(v)}\right) \\ + \exp\left(-\frac{\Lambda(v)t}{16A}\right)\left\{1 + \exp\left(A^{-1}\|x|_{d'}\|^{\star}_{\infty,v}\right)\right\}. \tag{3.4.20}$$

From certain optimality properties of rates of convergence obtained in Banerjee and Budhiraja (2020) (see Banerjee and Budhiraja (2020, Section 8)), we take $v = v^*$ where $v_i^* = \left(a^{(d')}\right)^{-1}\sigma_i$, $1 \leq i \leq d'$. Noting that $T(v^*) = T^{(d')}$, $\Lambda(v^*) = \Lambda^{(d')}$ and $\|x|_{d'}\|^{\star}_{\infty,v^*} \leq \|x|_{d'}\|_\infty/(\underline{\sigma}a^{(d')})$, the lemma follows from (3.4.20). $\square$

The following lemma combines the local contraction estimates obtained in Lemma 3.4.3 and the probability estimates on number of times subsets of coordinates hit zero by time t, obtained in Lemma 3.4.5, to furnish upper bounds on $\mathbb{E}\left[u(x,t)\right]$, $x \in \mathbb{R}_+^d, t \geq 0$.

Lemma 3.4.6. *Suppose Assumption 3.2.1 holds. Fix $d' \in \{k_0, \ldots, d\}$ and $x \in \mathbb{R}_+^d$. Recall $u(x,\cdot)$ from (3.2.1), the quantities $\lambda, \beta, \delta, C'$ in Lemma 3.4.3, and $A_0, \Lambda^{(d')}, T^{(d')}, \delta', C''$ in Lemma 3.4.5. Define*

$$\lambda(t) = \lambda^{\lfloor t\delta'/(4T^{(d')})\rfloor}. \tag{3.4.21}$$

Then for any $A \geq A_0$ and $t \geq 4T^{(d')}/\delta'$,

$$\mathbb{E}\left[u(x,t)\right] \leq u(x,0)\left[\exp\left(-t\frac{\delta' C''}{T^{(d')}}\right) + \exp\left(-t\frac{C''\Lambda^{(d')}}{A}\right)\left\{1 + \exp\left(\frac{\|x|_{d'}\|_\infty}{Aa^{(d')}\underline{\sigma}}\right)\right\}\right] \\ + u(x,0)\lambda(t) + C'\|x\|_{\infty,\delta}\left(\beta/\delta\right)^{d'+1}. \tag{3.4.22}$$

In the case $d' = d$, (3.4.22) holds without the $C'\|x\|_{\infty,\delta}\left(\beta/\delta\right)^{d'+1}$ term in the bound.

Proof. With $\tau(x, d')$ as in Corollary 3.4.4, we have for any $A \geq A_0$ and $t \geq 4T^{(d')}/\delta'$,

$$\begin{aligned}
\mathbb{E}\left[u(x,t)\right] &\leq \mathbb{E}\left[u(x,t)\mathbb{1}_{\tau(x,d')>t}\right] + C'\|x\|_{\infty,\delta}\left(\beta/\delta\right)^{d'+1} \\
&= \mathbb{E}\left[u(x,t)\mathbb{1}_{\tau(x,d')>t,\, \mathcal{N}_{d'}(x,t)<t\delta'/4T^{(d')}}\right] \\
&\quad + \mathbb{E}\left[u(x,t)\mathbb{1}_{\tau(x,d')>t,\, \mathcal{N}_{d'}(x,t)\geq t\delta'/4T^{(d')}}\right] + C'\|x\|_{\infty,\delta}\left(\beta/\delta\right)^{d'+1} \\
&\leq u(x,0)\,\mathbb{P}\left[\mathcal{N}_{d'}(x,t) < \delta' t/(4T^{(d')})\right] + \lambda(t)u(x,0) + C'\|x\|_{\infty,\delta}\left(\beta/\delta\right)^{d'+1} \\
&\leq u(x,0)\left[\exp\left(-t\frac{\delta' C''}{T^{(d')}}\right) + \exp\left(-t\frac{C''\Lambda^{(d')}}{A}\right)\left\{1 + \exp\left(\frac{\|x|_{d'}\|_\infty}{Aa^{(d')}\underline{\sigma}}\right)\right\}\right] \\
&\quad + \lambda(t)u(x,0) + C'\|x\|_{\infty,\delta}\left(\beta/\delta\right)^{d'+1}, \qquad (3.4.23)
\end{aligned}$$

where the second inequality follows from the monotonicity of u and Corollary 3.4.4, and the last inequality follows from Lemma 3.4.5.

When $d' = d$, by Corollary 3.4.4,

$$u(x,t)\ \mathbb{1}_{N_d(x,t)\geq t\delta'/4T^{(d)}} \leq \lambda(t)u(x,0).$$

Thus, again using the monotonicity of u and applying Lemma 3.4.5 with $d' = d$, we have for any $A \geq A_0$ and $t \geq 4T^{(d')}/\delta'$,

$$\begin{aligned}
&\mathbb{E}\left[u(x,t)\right] \leq \mathbb{E}\left[u(x,t)\ \mathbb{1}_{N_d(x,t)<t\delta'/4T^{(d)}}\right] + \mathbb{E}\left[u(x,t)\ \mathbb{1}_{N_d(x,t)\geq t\delta'/4T^{(d)}}\right] \\
&\leq u(x,0)\left[\exp\left(-t\frac{\delta' C''}{T^{(d)}}\right) + \exp\left(-t\frac{C''\Lambda^{(d)}}{A}\right)\left\{1 + \exp\left(\frac{\|x\|_\infty}{Aa^{(d)}\underline{\sigma}}\right)\right\}\right] + \lambda(t)u(x,0). \quad (3.4.24)
\end{aligned}$$

The lemma follows from (3.4.23) and (3.4.24). □

For any $x \in \mathbb{R}^d_+$ and $d' \in \{k_0, \ldots, d-1\}$, Lemma 3.4.6 shows that one can track the number of times the first d' co-ordinates of $X(x, \cdot)$ hit zero by time t to achieve exponential contraction in time t of the weighted distance $u(x, \cdot)$ between $X(x, \cdot)$ and $X(0, \cdot)$, till $u(x, \cdot)$ hits $C'\|x\|_{\infty,\delta}\,(\beta/\delta)^{d'+1}$. Thus, to ensure that this exponential contraction holds till $u(x, \cdot)$ is small, d' should be close to d. However, for large d, choosing a large d' slows down the convergence rate as it takes a long time for the d' co-ordinates to hit zero. This is manifested in the large value of $T^{(d')}$ which makes the exponential contraction coefficient in (3.4.22) small. In the next lemma, we take an *adaptive approach where the number of co-ordinates tracked increases with time.* Suppose Assumption 3.2.1 holds. With $r^* \geq 0$ as in III of Assumption 3.2.1, set

$$\ell(t) = \begin{cases} d \wedge \lfloor t^{1/(3+2r^*)} \rfloor & \text{under Assumption 3.2.1,} \\ d \wedge \lfloor t^{1/(1+2r^*)} \rfloor & \text{under Assumption 3.2.2.} \end{cases} \tag{3.4.25}$$

$\ell(\cdot)$ represents the time varying number of coordinates of the process $X(x, \cdot)$ that must hit zero to achieve a desired contraction. The choice of $\ell(\cdot)$ is obtained by optimizing bounds on the exponents appearing in (3.4.22) which depend on the assumptions.

Lemma 3.4.7 (Decay rate of $\mathbb{E}[u(x, \cdot)]$)**.** *Fix an initial condition $X(x, 0) = x \geq 0$. With δ, β as in Lemma 3.4.3, recall the weighted supremum norm $\|x\|_{\infty,\delta}$ and the process $u(x, \cdot)$ as in* (3.2.1)*. Define $\ell(\cdot)$ as in* (3.4.25)*.*

If Assumption 3.2.1 holds, there exist constants $C_0, C_1 > 0$ not depending on d, x, r^* such that, with $k_0' = k_0'(r^*) = k_0 \vee \left(\frac{8(3+2r^*)}{C_0' e}\right)^2$, we have for $d > k_0'$ and any $A \geq A_0$ (A_0 defined in Lemma 3.4.5),

$$\mathbb{E}[u(x,t)] \leq \begin{cases} C_1 \left(u(x,0) e^{\frac{\|x|_{\ell(t)}\|_\infty}{A \underline{\sigma} a^{(\ell(t))}}} + \|x\|_{\infty,\delta} \right) e^{-\frac{C_0}{A} t^{1/(3+2r^*)}} + C_1 u(x,0)\, e^{-C_0 \frac{t^{1/(3+2r^*)}}{\log t}}, & k_0' \leq \ell(t) < d, \\ C_1 u(x,0) e^{\frac{\|x\|_\infty}{A \underline{\sigma} a^{(d)}}} e^{-C_0 \frac{t}{A d^{2(1+r^*)}}} + C_1 u(x,0) e^{-C_0 \frac{t}{d^{2(1+r^*)} \log d}}, & \ell(t) = d. \end{cases} \tag{3.4.26}$$

If Assumption 3.2.2 holds, we have using the same constants k_0', C_0, C_1,

$$\mathbb{E}[u(x,t)] \leq \begin{cases} C_1 \left(u(x,0) e^{\frac{\|x|_{\ell(t)}\|_\infty}{A \underline{\sigma} a^{(\ell(t))}}} + \|x\|_{\infty,\delta} \right) e^{-\frac{C_0}{A} t^{1/(1+2r^*)}} + C_1 u(x,0)\, e^{-C_0 \frac{t^{1/(1+2r^*)}}{\log t}}, & k_0' \leq \ell(t) < d, \\ C_1 u(x,0) e^{\frac{\|x\|_\infty}{A \underline{\sigma} a^{(d)}}} e^{-C_0 \frac{t}{A d^{2r^*}}} + C_1 u(x,0) e^{-C_0 \frac{t}{d^{2r^*} \log d}}, & \ell(t) = d. \end{cases} \tag{3.4.27}$$

Proof. We will employ Lemma 3.4.6 with $d' = \ell(t)$. We will consider two cases: $k_0 \leq \ell(t) < d$ and $\ell(t) = d$.

In the work below, all constants depend on $\alpha, M, C, r^*, b_0, \underline{\sigma}, \overline{\sigma}$ in the notation of Assumptions 3.2.1 and 3.2.2, and $\beta \in (\alpha, 1)$ of Lemma 3.4.3.

Case ($k_0 \le \ell(t) < d$). *First suppose Assumption 3.2.1 holds. Set $d' = \ell(t)$ where for now we suppress the dependence on t. To employ the bound in Lemma 3.4.6, we consider bounds on the quantities $T^{(d')}, \Lambda^{(d')}$ and $a^{(d')}$. III of Assumption 3.2.1 implies $\underline{b}^{(d')} \ge b_0(d')^{-r^*}$ for some $b_0 > 0$ not depending on d. This along with II, IV of Assumption 3.2.1 gives*

$$a^{(d')} \le \overline{\sigma} \max_{1 \le i \le d'} \frac{1}{b_i^{(d')}} \sum_{j=1}^{d'} R_{ij}^{-1} \le \frac{d'\overline{\sigma}M}{\underline{b}^{(d')}} \le (d')^{1+r^*} \frac{\overline{\sigma}M}{b_0}. \tag{3.4.28}$$

Here, we have used $(R|_{d'})_{ij}^{-1} \le R_{ij}^{-1}$ for $1 \le i, j \le d'$ in the first inequality, which is a consequence of P^T having non-negative entries. From (3.4.28) *and the definitions in* (3.4.18)*, setting $A \ge A_0$ and recalling $d' = \ell(t) = d \wedge \lfloor t^{1/(3+2r^*)} \rfloor$, there exists $C_0' > 0$ not dependent on d', d, r^* such that for all $t \ge 2$,*

$$\begin{aligned} -t\frac{\Lambda^{(d')}}{A} &= -t\frac{1}{A(a^{(d')})^2} \le -t\frac{b_0^2}{A(d')^{2(1+r^*)}\left(\overline{\sigma}M\right)^2} \le -\frac{C_0'}{A} t^{1/(3+2r^*)}, \\ -t\frac{\delta'}{T^{(d')}} &\le -t\frac{\delta'}{1 + (d')^{2(1+r^*)}\left(\frac{\overline{\sigma}M}{b_0}\right)^2 \log(2d')} \le -C_0' \frac{t^{1/(3+2r^*)}}{\log t}. \end{aligned} \tag{3.4.29}$$

C_0' in the above can be taken to be $\left(2\left(\frac{\overline{\sigma}M}{b_0}\right)^2 + 2\log 2\right)^{-1}$. Recalling $\lambda(t) = \lambda^{\lfloor t\delta'/4T^{(d')}\rfloor}$, (3.4.29) *also gives*

$$\lambda(t) \le \lambda^{\frac{C_0'}{4}\frac{t^{1/(3+2r^*)}}{\log t} - 1}. \tag{3.4.30}$$

In addition, (3.4.29) *implies $\frac{4T^{(d')}}{\delta'} \le \frac{4}{C_0'} t^{1-1/(3+2r^*)} \log t$. Since as a function of t, $t^{-1/2(3+2r^*)} \log t$ is upper-bounded by $\frac{2(3+2r^*)}{e}$, we have $\frac{4T^{(d')}}{\delta'} \le t$ for $\left(\frac{8(3+2r^*)}{C_0'e}\right)^{2(3+2r^*)} \le t$. These calculations show the condition $t \ge \frac{4T^{(d')}}{\delta'}$ in Lemma 3.4.6 holds when $\ell(t) \ge \left(\frac{8(3+2r^*)}{C_0'e}\right)^2$.*

We now apply (3.4.29), (3.4.30) *to* (3.4.22) *in Lemma 3.4.6, with $A \ge A_0$. Recalling $k_0' = k_0'(r^*) = k_0 \vee \left(\frac{8(3+2r^*)}{C_0'e}\right)^2$ we have*

$$\begin{aligned} \mathbb{E}\left[u(x,t)\right] \le u(x,0)\left[1 + \exp\left(\frac{\|x|_{\ell(t)}\|_\infty}{A\underline{\sigma}a^{(\ell(t))}}\right)\right] e^{-\frac{C''C_0'}{A}t^{1/(3+2r^*)}} + u(x,0)e^{-C''C_0'\frac{t^{1/(3+2r^*)}}{\log t}} \\ + u(x,0)\,\lambda^{\frac{C_0'}{4}\frac{t^{1/(3+2r^*)}}{\log t} - 1} + C'\|x\|_{\infty,\delta}\,(\beta/\delta)^{t^{1/(3+2r^*)}}, \qquad \text{for} \quad k_0' \le \ell(t) < d, \end{aligned} \tag{3.4.31}$$

where we used in the second line $d'+1=\ell(t)+1\geq t^{1/(3+2r^*)}$*. This proves the first case in* (3.4.26) *with*

$$\begin{aligned} C_0 &= C''C_0'\wedge\frac{C_0'}{4}\log\frac{1}{\lambda}\wedge\log\frac{\delta}{\beta}, \\ C_1 &= \left(2+\frac{1}{\lambda}\right)\vee C'. \end{aligned} \tag{3.4.32}$$

If Assumption 3.2.2 holds, we set $d'=\ell(t)=d\wedge\lfloor t^{1/(1+2r^*)}\rfloor$*. Instead of* (3.4.28) *we have*

$$a^{(d')}\leq\overline{\sigma}\max_{1\leq i\leq d'}\frac{1}{b_i^{(d')}}\sum_{j=1}^{d'}R_{ij}^{-1}\leq\frac{\overline{\sigma}M}{\underline{b}^{(d')}}\leq(d')^{r^*}\frac{\overline{\sigma}M}{b_0}. \tag{3.4.33}$$

Proceeding in the same way as (3.4.29)*, we use* (3.4.33) *to show*

$$\begin{aligned} -t\frac{\Lambda^{(d')}}{A} &= -t\frac{1}{A(a^{(d')})^2}\leq-t\frac{b_0^2}{A(d')^{2r^*}\left(\overline{\sigma}M\right)^2}\leq-\frac{C_0'}{A}t^{1/(1+2r^*)}, \\ -t\frac{\delta'}{T^{(d')}} &\leq -t\frac{\delta'}{1+(d')^{2r^*}\left(\frac{\overline{\sigma}M}{b_0}\right)^2\log\left(2d'\right)}\leq-C_0'\frac{t^{1/(1+2r^*)}}{\log t}. \end{aligned} \tag{3.4.34}$$

Arguing as in (3.4.30) *but using* (3.4.34) *instead of* (3.4.29)*, we have*

$$\lambda(t)\leq\lambda^{\frac{C_0'}{4}\frac{t^{1/(1+2r^*)}}{\log t}-1} \tag{3.4.35}$$

Using (3.4.34)*, we have* $\frac{4T^{(d')}}{\delta'}\leq\frac{4}{C_0'}t^{1-1/(1+2r^*)}\log t$*. As in the argument after* (3.4.30)*,* $\frac{4T^{(d')}}{\delta'}\leq t$ *for* t *such that* $\ell(t)\geq\left(\frac{8(1+2r^*)}{C_0'e}\right)^2$*, under which Lemma 3.4.6 is valid. We now apply* (3.4.34)*,* (3.4.35) *to* (3.4.22)*:*

$$\begin{aligned} \mathbb{E}\left[u(x,t)\right]\leq u(x,0)\left[1+\exp\left(\frac{\|x|_{\ell(t)}\|_\infty}{Aa^{(\ell(t))}\underline{\sigma}}\right)\right]e^{-\frac{C''C_0'}{A}t^{1/(1+2r^*)}}+u(x,0)e^{-C''C_0'\frac{t^{1/(1+2r^*)}}{\log t}} \\ +u(x,0)\,\lambda^{\frac{C_0'}{4}\frac{t^{1/(1+2r^*)}}{\log t}-1}+C'\|x\|_{\infty,\delta}\left(\beta/\delta\right)^{t^{1/(1+2r^*)}},\qquad\textit{for}\quad k_0'\leq\ell(t)<d. \end{aligned} \tag{3.4.36}$$

This proves the first case in (3.4.27) *with* C_0,C_1 *as in* (3.4.32)*. Since* $\left(\frac{8(1+2r^*)}{C_0'e}\right)^2<\left(\frac{8(3+2r^*)}{C_0'e}\right)^2$*, we use the same* k_0' *in* (3.4.31) *and* (3.4.36)*.*

Case $(\ell(t) = d)$**.** *First we consider the scenario of Assumption 3.2.1, in which case* $\ell(t) = d$ *implies* $t \geq d^{3+2r^*}$. *We follow the same basic recipe: We use Lemma 3.4.6, this time in the case* $d' = d$*, and bound the quantities* $a^{(d)}, T^{(d)}, \Lambda^{(d)}$.

The bound on $a^{(d')}$ *in* (3.4.28) *continues to hold with* $d' = d$*, and using this with we have*

$$\begin{aligned} -t\frac{\Lambda^{(d)}}{A} &= -t\frac{1}{A(a^{(d)})^2} \leq -t\frac{Ab_0^2}{d^{2(1+r^*)}\left(\overline{\sigma}M\right)^2} \leq -C_0'\frac{t}{d^{2(1+r^*)}}, \\ -t\frac{\delta'}{T^{(d)}} &\leq -t\frac{\delta'}{1+d^{2(1+r^*)}\left(\frac{\overline{\sigma}M}{b_0}\right)^2\log(2d)} \leq -C_0'\frac{t}{d^{2(1+r^*)}\log d}. \end{aligned} \tag{3.4.37}$$

Now (3.4.37) *implies*

$$\lambda(t) \leq \lambda^{\frac{C_0'}{4}\frac{t}{d^{2(1+r^*)}\log d}-1}. \tag{3.4.38}$$

Since the lemma statement has imposed $d > k_0'$ *we have* $t \geq [k_0'(r^*)]^{3+2r^*}$. *Applying the argument preceding* (3.4.31)*, this implies* $t \geq \frac{4T^{(d)}}{\delta'}$ *and thus Lemma 3.4.6 holds for all* $t \geq [k_0'(r^*)]^{3+2r^*}$ *in the case* $\ell(t) = d$. *Using* (3.4.37), (3.4.38) *in* (3.4.22) *(without the* $C'\|x\|_{\infty,\delta}\left(\beta/\delta\right)^{d'+1}$ *term) we have*

$$\mathbb{E}\left[u(x,t)\right] \leq C_1 u(x,0) e^{\frac{\|x\|_\infty}{Aa^{(d)}\underline{\sigma}}} e^{-C_0\frac{t}{Ad^{2(1+r^*)}}} + C_1 u(x,0) e^{-C_0\frac{t}{d^{2(1+r^*)}\log d}}, \qquad \textit{for} \quad \ell(t) = d, \tag{3.4.39}$$

where we use C_0, C_1 *from* (3.4.32). *This is the second line in* (3.4.26).

When Assumption 3.2.2 holds, $\ell(t) = d$ *implies* $t \geq d^{1+2r^*}$. *The second line of* (3.4.27) *is proven in identical fashion to* (3.4.39)*, after accounting for the stronger assumptions in the same way as we did in* (3.4.34) *and* (3.4.36).

□

For any $\beta \in (\alpha, 1)$ and $x \in \mathbb{R}_+^d$, Lemma 3.4.7 gives quantitative estimates for the decay rate of the $\|\cdot\|_{1,\beta}$ distance between $X(x,\cdot)$ and $X(0,\cdot)$. To use this in furnishing rates of convergence to stationarity in $\|\cdot\|_{1,\beta}$ distance starting from any $x \in \mathbb{R}_+^d$, namely Theorem 3.2.1, we use Theorem 3.4.1 to make the following simple observation. Recalling u in (3.2.1) and u_π in (3.2.2), we have by the triangle inequality,

$$\begin{aligned}\| (X(x,t) - X(X(\infty),t)) \|_{1,\beta} &\leq \| (X(x,t) - X(0,t)) \|_{1,\beta} + \| (X(X(\infty),t) - X(0,t)) \|_{1,\beta} \\ &\leq \|R^{-1} (X(x,t) - X(0,t)) \|_{1,\beta} \\ &+ \|R^{-1} (X(X(\infty),t) - X(0,t)) \|_{1,\beta} \\ &= u(x,t) + u_\pi(t). \end{aligned} \tag{3.4.40}$$

To bound the expectation of the final two terms in (3.4.40), we apply Lemma 3.4.7 to bound $\mathbb{E}[u(x,t)]$. To bound $\mathbb{E}[u_\pi(t)]$, we will use a slightly altered version of Lemma 3.4.6 and Lemma 3.4.7 conditional on $x = X(\infty)$ followed by taking expectation in the law of $X(\infty)$. This will require quantitative control over moments of several functionals of $X(\infty)$. This is the objective of the following lemma.

Lemma 3.4.8 (Moments under stationarity)**.** *Suppose Assumption 3.2.1 holds, with $\alpha \in (0,1)$ set therein. Fix $\beta \in (\alpha, 1)$ and define $u(x,0) = \|x\|_{1,\beta}$ as in* (3.2.1)*. Fix $\delta \in (\beta, 1)$.*

Recall the random variable $X(\infty)$ distributed as the stationary distribution for the process (3.1.1)*. Fix $d' \in \{k_0 \dots d\}$. Then there exists a constant $C''' \geq 1$ not depending on d', d or r^* (see III of Assumption 3.2.1) such that*

$$\begin{aligned} \mathbb{E}\left[\exp\left(2\frac{\|X|_{d'}(\infty)\|_\infty}{A\underline{\sigma}a^{(d')}}\right)\right] &\leq 1 + d' \qquad \textit{for } A \geq 2d'\frac{\overline{\sigma}M}{\underline{\sigma}}, && (3.4.41) \\ \mathbb{E}\left[\|X(\infty)\|_{\infty,\delta}\right] \leq \mathbb{E}\left[\|X(\infty)\|_{\infty,\sqrt{\delta}}\right] &\leq C''' L_1(\delta), && (3.4.42) \\ \mathbb{E}\left[u\left(X(\infty),0\right)\right] \leq \sqrt{\mathbb{E}\left[u^2\left(X(\infty),0\right)\right]} &\leq C''' L_1(\delta). && (3.4.43) \end{aligned}$$

where $L_1(\delta) := \left(k_0^{r^*+1} + \sum_{i=k_0}^d i^{3+r^*}\delta^{i/2}\right)$.

If in addition Assumption 3.2.2 holds, we have

$$\begin{aligned} \mathbb{E}\left[\exp\left(2\frac{\|X|_{d'}(\infty)\|_\infty}{A\underline{\sigma}a^{(d')}}\right)\right] &\leq 1 + d' \qquad \textit{for } A \geq 2\frac{\overline{\sigma}M}{\underline{\sigma}}, && (3.4.44) \\ \mathbb{E}\left[\|X(\infty)\|_{\infty,\delta}\right] \leq \mathbb{E}\left[\|X(\infty)\|_{\infty,\sqrt{\delta}}\right] &\leq C''' L_2(\delta), && (3.4.45) \\ \mathbb{E}\left[u\left(X(\infty),0\right)\right] \leq \sqrt{\mathbb{E}\left[u^2\left(X(\infty),0\right)\right]} &\leq C''' L_2(\delta). && (3.4.46) \end{aligned}$$

where $L_2(\delta) := \left(k_0^{r^*} + \sum_{i=k_0}^{d} i^{2+r^*}\delta^{i/2}\right)$.

Proof. For $k \in \{k_0, \ldots, d\}$, write $\bar{X}^{(k)}$ for the process $\bar{X}$ defined in Theorem 3.4.2.

III of Assumption 3.2.1 imposes $-\left(R|_{d'}\right)^{-1}\mu|_{d'} = b^{(d')} > 0$. Thus $\bar{X}^{(d')}$ has a stationary distribution (Harrison and Williams (1987a), Section 6). We write $\bar{X}^{(d')}(\infty)$ for the random variable with this distribution. From Harrison and Williams (1987a, Lemma 4, Section 6) and Harrison and Williams (1987a, Lemma 12 and its proof, Section 6), for any $\theta^{(d')} \in \mathbb{R}^{d'}$ such that

$$\theta^{(d')} > 0, \qquad \left(R|_{d'}\right)^{-1}\theta^{(d')} \leq b^{(d')}, \tag{3.4.47}$$

we have

$$\begin{aligned}\mathbb{P}\left[\left(R|_{d'}\right)^{-1}\bar{X}^{(d')}(\infty) \leq \left(R|_{d'}\right)^{-1} z\right] &= \liminf_{t\to\infty}\mathbb{P}\left[\left(R|_{d'}\right)^{-1}\bar{X}^{(d')}(0,t) \leq \left(R|_{d'}\right)^{-1} z\right]\\ &\geq 1 - \sum_{j=1}^{d'}\exp\left(-2z_j\frac{\theta_j^{(d')}}{\sigma_j}\right), \qquad z \in \mathbb{R}_+^{d'}.\end{aligned} \tag{3.4.48}$$

In other words, the distribution of $\left(R|_{d'}\right)^{-1}\bar{X}^{(d')}(\infty)$ has exponential tails. This is the key fact in proving the lemma, and the remainder of the argument is in choosing $\theta^{(d')}$ appropriately to achieve the desired dependence on the parameters and dimension. Recalling the quantity $a^{(d')}$ from Assumption 3.2.1 we set

$$\theta_i^{(d')} = \frac{\underline{\sigma}}{a^{(d')}}, \qquad 1 \leq i \leq d'. \tag{3.4.49}$$

By definition of $a^{(d')}$, for each $1 \leq i \leq d'$,

$$\begin{aligned}\sum_{\ell=1}^{d'}\left(R|_{d'}\right)_{i\ell}^{-1}\theta_\ell^{(d')} &\leq \sum_{\ell=1}^{d'}\left(R|_{d'}\right)_{i\ell}^{-1}\left(\frac{b_i^{(d')}\underline{\sigma}}{\sum_{j=1}^{d'}\left(R|_{d'}\right)_{ij}^{-1}\sigma_j}\right)\\ &\leq \sum_{\ell=1}^{d'}\left(R|_{d'}\right)_{i\ell}^{-1}\left(\frac{b_i^{(d')}\sigma_\ell}{\sum_{j=1}^{d'}\left(R|_{d'}\right)_{ij}^{-1}\sigma_j}\right) = b_i^{(d')},\end{aligned}$$

and hence, $\theta^{(d')}$ satisfies (3.4.47).

We now prove the exponential moments (3.4.41) and (3.4.44). Since we consider a fixed d' here, we write $\bar{X}(\infty) = \bar{X}^{(d')}(\infty)$ to lighten notation. Note Theorem 3.4.2 implies $\bar{X}(\infty)$

stochastically dominates $X|_{d'}(\infty)$. Hence, since $(R|_{d'})^{-1}_{ij} \geq 0$ we have for any $z \in \mathbb{R}^{d'}_+$,

$$\mathbb{P}\left[\left((R|_{d'})^{-1} X|_{d'}(\infty)\right)_i \leq z_i\,, \quad 1 \leq i \leq d'\right] \geq \mathbb{P}\left[\left((R|_{d'})^{-1} \bar{X}(\infty)\right)_i \leq z_i\,, \quad 1 \leq i \leq d'\right]. \tag{3.4.50}$$

For arbitrary $z_0 \geq 1$, setting $z_i = (\log z_0)\frac{Aa^{(d')}}{2}$ for each $i = 1, \ldots, d'$ in (3.4.50),

$$\begin{aligned}
&\mathbb{P}\left[\exp\left(2\frac{\|X|_{d'}(\infty)\|_\infty}{Aa^{(d')}}\right) \leq z_0\right] = \mathbb{P}\left[X_i(\infty) \leq (\log z_0)\frac{Aa^{(d')}}{2}\,, \quad 1 \leq i \leq d'\right] \\
&\geq \mathbb{P}\left[\left((R|_{d'})^{-1} X|_{d'}(\infty)\right)_i \leq (\log z_0)\frac{Aa^{(d')}}{2}\,, \quad 1 \leq i \leq d'\right] \\
&\geq \mathbb{P}\left[\left((R|_{d'})^{-1} X|_{d'}(\infty)\right)_i \leq (\log z_0)\frac{Aa^{(d')}}{2}\frac{\sum_{j=1}^{d'}(R|_{d'})^{-1}_{ij}}{\max_{1\leq k\leq d'}\sum_{j=1}^{d'}(R|_{d'})^{-1}_{kj}}\,, \quad 1 \leq i \leq d'\right] \\
&\geq 1 - \sum_{j=1}^{d'} \exp\left(-(\log z_0)\frac{A\underline{\sigma}}{\sigma_j \max_{1\leq k\leq d'}\sum_{j=1}^{d'}(R|_{d'})^{-1}_{kj}}\right) \geq 1 - d' \exp\left(-(\log z_0)\frac{A\underline{\sigma}}{\overline{\sigma} M d'}\right).
\end{aligned} \tag{3.4.51}$$

We used in the second line $(R|_{d'})^{-1} x \geq x$, $\forall\, x \in \mathbb{R}^{d'}_+$. For the third inequality, we used (3.4.48) with the d'-dimensional vector $(\log z_0)\frac{Aa^{(d')}}{2}\left(\max_{1\leq k\leq d'}\sum_{j=1}^{d'}(R|_{d'})^{-1}_{kj}\right)^{-1}(1, \ldots, 1)^T$ in place of z. For the last inequality, we used II and IV of Assumption 3.2.1. From (3.4.51) we obtain (3.4.41) as follows:

$$\begin{aligned}
\mathbb{E}\left[\exp\left(2\frac{\|X|_{d'}(\infty)\|_\infty}{A\underline{\sigma}a^{(d')}}\right)\right] &\leq 1 + \int_1^\infty \mathbb{P}\left[\exp\left(2\frac{\|X|_{d'}(\infty)\|_\infty}{Aa^{(d')}}\right) > z_0\right] dz_0 \\
&\leq 1 + d' \int_1^\infty \exp\left(-(\log z_0)\frac{A\underline{\sigma}}{\overline{\sigma} M d'}\right) \leq 1 + d', \qquad A \geq 2d'\frac{\overline{\sigma} M}{\underline{\sigma}}.
\end{aligned} \tag{3.4.52}$$

If instead Assumption 3.2.2 holds, then $\max_{1\le k\le d'}\sum_{j=1}^{d'}(R|_{d'})^{-1}_{kj}\le \max_{1\le k\le d'}\sum_{j=1}^{d'}R^{-1}_{kj}\le M$. Instead of (3.4.51) we have,

$$\begin{aligned}
&\mathbb{P}\left[\exp\left(2\frac{\|X|_{d'}(\infty)\|_\infty}{Aa^{(d')}}\right)\le z_0\right]=\mathbb{P}\left[X_i(\infty)\le(\log z_0)\frac{Aa^{(d')}}{2}\,,\quad 1\le i\le d'\right]\\
&\ge 1-\sum_{j=1}^{d'}\exp\left(-(\log z_0)\frac{A\underline{\sigma}}{\sigma_j\max_{1\le k\le d'}\sum_{j=1}^{d'}(R|_{d'})^{-1}_{kj}}\right)\\
&\ge 1-\sum_{j=1}^{d'}\exp\left(-(\log z_0)\frac{A\underline{\sigma}}{\sigma_j M}\right)\ge 1-d'\exp\left(-(\log z_0)\frac{A\underline{\sigma}}{\overline{\sigma}M}\right). \qquad (3.4.53)
\end{aligned}$$

This proves (3.4.44) by proceeding exactly as in (3.4.52), using (3.4.53) in place of (3.4.51).

We turn to (3.4.42), recalling the notation $\bar{X}^{(k)}(\infty)$ from the start of this proof. By Theorem 3.4.2, $X_i(\infty)\le\bar{X}^{(k_0)}_i(\infty)$ for $i=1,\dots,k_0$ and $X_i(\infty)\le\bar{X}^{(i)}_i(\infty)$ for $i=k_0,\dots,d$. This implies

$$\mathbb{P}\left[\|X(\infty)\|_{\infty,\sqrt{\delta}}>z_0\right]\le\mathbb{P}\left[\|\bar{X}^{(k_0)}(\infty)\|_{\infty,\sqrt{\delta}}>z_0\right]+\sum_{i=k_0}^{d}\mathbb{P}\left[\delta^{i/2}\bar{X}^{(i)}_i(\infty)>z_0\right]. \qquad (3.4.54)$$

In preparation to handle the first probability of the right-hand side in (3.4.54), we note that by I, II of Assumption 3.2.1 and $(R|_{k_0})^{-1}_{ij}\le R^{-1}_{ij}$,

$$\begin{aligned}
\sum_{j=1}^{k_0}(R|_{k_0})^{-1}_{ij}\,\delta^{-j/2}&\le M\sum_{j=1}^{i}\delta^{-j/2}+\mathbb{1}_{i<k_0}C\sum_{j=i+1}^{k_0}\alpha^{j-i}\delta^{-j/2}\\
&\le\delta^{-i/2}\frac{M}{1-\sqrt{\delta}}+\mathbb{1}_{i<k_0}C\delta^{-i/2}\sum_{j=1}^{k_0-i}\left(\alpha/\sqrt{\delta}\right)^j\le C'\delta^{-i/2}, \qquad (3.4.55)
\end{aligned}$$

for $C'=\frac{M}{1-\sqrt{\delta}}+C\frac{\alpha/\sqrt{\delta}}{1-(\alpha/\sqrt{\delta})}$, recalling that $0<\alpha<\sqrt{\delta}<1$. In the following, we set $\theta^{(i)}_j=\frac{\underline{\sigma}}{a^{(k_0)}}$ for $k_0\le i\le d, 1\le j\le i$. Note that,

$$\begin{aligned}
&\mathbb{P}\left[\|\bar{X}^{(k_0)}(\infty)\|_{\infty,\sqrt{\delta}}\le z_0\right]\ge\mathbb{P}\left[\left((R|_{k_0})^{-1}\bar{X}^{(k_0)}(\infty)\right)_i\le z_0\delta^{-i/2},\quad i=1,\dots,k_0\right]\\
&\ge\mathbb{P}\left[\left((R|_{k_0})^{-1}\bar{X}^{(k_0)}(\infty)\right)_i\le z_0\frac{\sum_{j=1}^{k_0}(R|_{k_0})^{-1}_{ij}\,\delta^{-j/2}}{C'},\quad i=1,\dots,k_0\right]\\
&\ge 1-\sum_{j=1}^{k_0}\exp\left(-2z_0\delta^{-j/2}\frac{\theta^{(k_0)}_j}{C'\sigma_j}\right)\ge 1-\sum_{j=1}^{k_0}\exp\left(-2z_0\delta^{-j/2}\frac{\underline{\sigma}}{a^{(k_0)}C'\overline{\sigma}}\right). \qquad (3.4.56)
\end{aligned}$$

In the first line we used $\left((R|_{k_0})^{-1}\bar{X}_i^{(k_0)}(\infty)\right)_i \geq \bar{X}_i^{(k_0)}(\infty),\ i=1,\ldots,k_0$. The second line uses (3.4.55). The last line applies (3.4.48) with k_0 in place of d' and with $z_j=(C')^{-1}z_0\delta^{-j/2},\ 1\leq j\leq k_0$, and uses IV of Assumption 3.2.1.

Now we bound $\mathbb{P}\left[\delta^{i/2}\bar{X}_i^{(i)}(\infty)>z_0\right]$ for $i=k_0,\ldots,d$ required to bound the second term of the right hand side in (3.4.54). In the following equations we use II of Assumption 3.2.1, which says $(R|_i)^{-1}_{kj}\leq M$, in the third line and (3.4.48) to show

$$
\begin{aligned}
\mathbb{P}\left[\delta^{i/2}\bar{X}_i^{(i)}(\infty)\leq z_0\right] &\geq \mathbb{P}\left[\left((R|_i)^{-1}\bar{X}^{(i)}(\infty)\right)_i\leq z_0\delta^{-i/2}\right]\\
&\geq \mathbb{P}\left[\left((R|_i)^{-1}\bar{X}^{(i)}(\infty)\right)_k\leq z_0\delta^{-i/2},\quad k=1,\ldots,i\right]\\
&\geq \mathbb{P}\left[\left((R|_i)^{-1}\bar{X}^{(i)}(\infty)\right)_k\leq z_0\delta^{-i/2}\frac{\sum_{j=1}^{i}(R|_i)^{-1}_{kj}}{iM},\quad k=1,\ldots,i\right]\\
&\geq 1-\sum_{j=1}^{i}\exp\left(-2z_0\delta^{-i/2}\frac{\theta_j^{(i)}}{iM\sigma_j}\right)\geq 1-i\exp\left(-2z_0\delta^{-i/2}\frac{\underline{\sigma}}{ia^{(i)}M\overline{\sigma}}\right),
\end{aligned}
\tag{3.4.57}
$$

for $i=k_0,\ldots,d$, where in the first line we used $\left((R|_{k_0})^{-1}\bar{X}^{(i)}(\infty)\right)_i\geq\bar{X}_i^{(i)}(\infty)$. Applying (3.4.56), (3.4.57) to (3.4.54) we get

$$
\mathbb{P}\left[\|X(\infty)\|_{\infty,\sqrt{\delta}}>z_0\right]\ \leq\ \sum_{i=1}^{k_0}\exp\left(-2z_0\delta^{-i/2}\frac{\underline{\sigma}}{a^{(k_0)}C'\overline{\sigma}}\right)+\sum_{i=k_0}^{d}i\exp\left(-2z_0\delta^{-i/2}\frac{\underline{\sigma}}{ia^{(i)}M\overline{\sigma}}\right).
\tag{3.4.58}
$$

As a result,

$$
\begin{aligned}
&\mathbb{E}\left[\|X(\infty)\|_{\infty,\sqrt{\delta}}\right]=\int_0^\infty\mathbb{P}\left[\|X(\infty)\|_{\infty,\sqrt{\delta}}>z_0\right]\,dz_0\\
&\leq a^{(k_0)}\frac{\overline{\sigma}C'}{2\underline{\sigma}}\sum_{i=1}^{k_0}\delta^{i/2}+\frac{M\overline{\sigma}}{2\underline{\sigma}}\sum_{i=k_0}^{d}i^2a^{(i)}\delta^{i/2}\leq a^{(k_0)}\frac{\overline{\sigma}C'\sqrt{\delta}}{2\underline{\sigma}(1-\sqrt{\delta})}+\frac{M\overline{\sigma}}{2\underline{\sigma}}\sum_{i=k_0}^{d}i^2a^{(i)}\delta^{i/2}\\
&\leq k_0^{r^*+1}\frac{MC'\overline{\sigma}^2\sqrt{\delta}}{2b_0\underline{\sigma}(1-\sqrt{\delta})}+\frac{M^2\overline{\sigma}^2}{2b_0\underline{\sigma}}\sum_{i=k_0}^{d}i^{3+r^*}\delta^{i/2}\leq C''\left(k_0^{r^*+1}+\sum_{i=k_0}^{d}i^{3+r^*}\delta^{i/2}\right),
\end{aligned}
\tag{3.4.59}
$$

with $C''=\frac{M^2\overline{\sigma}^2}{2b_0\underline{\sigma}}\vee\frac{MC'\overline{\sigma}^2\sqrt{\delta}}{2b_0\underline{\sigma}(1-\sqrt{\delta})}$. In the final line we used the fact that $a^{(i)}\leq\frac{iM\overline{\sigma}}{\underline{b}^{(i)}}$ by definition of $a^{(i)}$ and Assumption 3.2.1, and $\underline{b}^{(i)}\geq b_0i^{-r^*}$ using III of Assumption 3.2.1. If instead

Assumption 3.2.2 holds, $a^{(i)} \le \frac{M\overline{\sigma}}{\underline{b}^{(i)}} \le i^{r^*}\frac{M\overline{\sigma}}{b_0}$. Substituting this fact in the final line of (3.4.59), but otherwise proceeding in exactly the same way, produces (3.4.45) with the same choice of C''.

Now we show (3.4.43) and (3.4.46). We need prove only the second inequality in (3.4.43), (3.4.46). Using Jensen's inequality in the first line below, we have

$$
\begin{aligned}
u^2(X(\infty),0) = \left(\sum_{i=1}^{d}\beta^i\sum_{j=1}^{d}R^{-1}_{ij}(\infty)X_j(\infty)\right)^2 &\le \frac{\beta}{1-\beta}\sum_{i=1}^{d}\beta^i\left(\sum_{j=1}^{d}R^{-1}_{ij}(\infty)X_j(\infty)\right)^2 \\
&\le \|X(\infty)\|^2_{\infty,\sqrt{\delta}}\frac{\beta}{1-\beta}\sum_{i=1}^{d}\beta^i\left(\sum_{j=1}^{d}R^{-1}_{ij}(\infty)\delta^{-j/2}\right)^2 \\
&\le \|X(\infty)\|^2_{\infty,\sqrt{\delta}}\frac{(C')^2\beta}{1-\beta}\sum_{i=1}^{d}\beta^i\delta^{-i} \\
&\le \|X(\infty)\|^2_{\infty,\sqrt{\delta}}\frac{(C')^2\beta}{1-\beta}\frac{\beta/\delta}{1-\beta/\delta}.
\end{aligned}
\tag{3.4.60}
$$

In the second line, we used (3.4.55) with d in place of k_0 and C' set therein. For the final line recall $\beta \in (\alpha,\delta)$. Using (3.4.58) to bound the quadratic moment on the right-hand side of (3.4.60),

$$
\begin{aligned}
&\mathbb{E}\left[u^2(X(\infty),0)\right] \le \frac{(C')^2\beta}{1-\beta}\frac{\beta/\delta}{1-\beta/\delta}\ \mathbb{E}\left[\|X(\infty)\|^2_{\infty,\sqrt{\delta}}\right] \\
&= \frac{(C')^2\beta}{1-\beta}\frac{\beta/\delta}{1-\beta/\delta}\int_0^\infty \mathbb{P}\left[\|X(\infty)\|_{\infty,\sqrt{\delta}} > \sqrt{z_0}\right]\,dz_0 \\
&\le \frac{(C')^2\beta}{1-\beta}\frac{\beta/\delta}{1-\beta/\delta}\int_0^\infty\left(\sum_{i=1}^{k_0}\exp\left(-2\sqrt{z_0}\delta^{-i/2}\frac{\underline{\sigma}}{a^{(k_0)}C'\overline{\sigma}}\right)+\sum_{i=k_0}^{d} i\exp\left(-2\sqrt{z_0}\delta^{-i/2}\frac{\underline{\sigma}}{ia^{(i)}M\overline{\sigma}}\right)\right)dz_0 \\
&= 2\frac{(C')^2\beta}{1-\beta}\frac{\beta/\delta}{1-\beta/\delta}\left(\left(\frac{a^{(k_0)}C'\overline{\sigma}}{2\underline{\sigma}}\right)^2\sum_{i=1}^{k_0}\delta^i+\left(\frac{M\overline{\sigma}}{2\underline{\sigma}}\right)^2\sum_{i=k_0}^{d}\delta^i i^3\left(a^{(i)}\right)^2\right).
\end{aligned}
\tag{3.4.61}
$$

Under Assumption 3.2.1 we have $a^{(i)} \le \frac{iM\overline{\sigma}}{\underline{b}^{(i)}} \le i^{1+r^*}\frac{M\overline{\sigma}}{b_0}$, and applying this to (3.4.61) gives

$$
\mathbb{E}\left[u^2(X(\infty),0)\right] \le (C''')^2\left(k_0^{2(r^*+1)}+\sum_{i=k_0}^{d} i^{5+2r^*}\delta^i\right),
\tag{3.4.62}
$$

where we have chosen $C''' \geq C''$ to be large enough that both (3.4.62) and (3.4.59) are satisfied,

$$C''' = C'' \vee \sqrt{\frac{(C')^2\beta}{2(1-\beta)}\frac{\beta/\delta}{1-\beta/\delta}\left(\frac{M\overline{\sigma}}{b_0}\right)^2\left(\left(\frac{C'\overline{\sigma}}{\underline{\sigma}}\right)^2\frac{\delta}{1-\delta}+\left(\frac{M\overline{\sigma}}{\underline{\sigma}}\right)^2\right)}. \tag{3.4.63}$$

Under Assumption 3.2.2 we have $a^{(i)} \leq \frac{M\overline{\sigma}}{\underline{b}^{(i)}} \leq i^{r^*}\frac{M\overline{\sigma}}{b_0}$, so by (3.4.61)

$$\mathbb{E}\left[u^2(X(\infty),0)\right] \leq (C''')^2\left(k_0^{2r^*}+\sum_{i=k_0}^{d} i^{3+2r^*}\delta^i\right). \tag{3.4.64}$$

After taking square roots and using $\sum_1^m x_i^2 \leq \left(\sum_1^m x_i\right)^2$ for any non-negative numbers $x_1 \dots x_m$, (3.4.62) proves (3.4.43) and (3.4.64) proves (3.4.46). $\square$

Now we bound $\mathbb{E}\left[u_\pi(t)\right]$. We would like simply to use Lemma 3.4.7 conditional on $x = X(\infty)$ followed by taking expectation in the law of $X(\infty)$. We will do so to prove (3.4.67) under Assumption 3.2.2, but this is not desirable under Assumption 3.2.1 for the following reason.

If one tries this approach under Assumption 3.2.1, terms of the form $\mathbb{E}\left[\exp\left(2\frac{\|X|_{\ell(t)}(\infty)\|_\infty}{A\underline{\sigma}a^{(\ell(t))}}\right)\right]$ (where $\ell(\cdot)$ is defined in (3.4.25)) appear in the bound and A should be chosen large enough so that this expectation is finite. Lemma 3.4.8 shows this requires A to be of order $\ell(t)$. However, such choice of A implies that $e^{-\frac{C_0}{A}t^{1/(3+2r^*)}}$ is bounded below by a positive dimension independent constant as $t \to \infty$, thereby lending the bounds obtained via Lemma 3.4.7 trivial.

Thus, under Assumption 3.2.1, we proceed by choosing a higher number of coordinates of $X(x,\cdot)$ that must hit zero in order to achieve a desirable contraction in $\mathbb{E}\left[u_\pi(\cdot)\right]$. Namely, instead of $\ell(\cdot)$ of Lemma 3.4.7, we define

$$d(t) = \begin{cases} d \wedge \lfloor t^{1/(4+2r^*)} \rfloor & \text{under Assumption 3.2.1,} \\ d \wedge \lfloor t^{1/(1+2r^*)} \rfloor & \text{under Assumption 3.2.2,} \end{cases} \tag{3.4.65}$$

with $r^* \geq 0$ as in III of Assumption 3.2.1.

Lemma 3.4.9 (Decay rate of $\mathbb{E}\left[u_\pi(\cdot)\right]$)**.** *Suppose Assumption 3.2.1 holds for X, an* RBM(Σ,μ,R)*, with $\alpha \in (0,1)$ defined therein. Fix $\beta \in (\alpha,1)$, $\delta \in (\beta,1)$, and recall the weighted distance $u_\pi(\cdot)$ from* (3.2.2)*. Recall $L_1(\delta), L_2(\delta)$ from Lemma 3.4.8.*

There exist constants $\bar{C}_0, \bar{C}_1, C_0' > 0$ not depending on d or r^* such that, with $k_0'' = k_0''(r^*) = \max\left\{k_0, \frac{A_0\underline{\sigma}}{2\overline{\sigma}M}, \left(\frac{8(4+2r^*)}{C_0' e}\right)^2\right\}$ (A_0 defined in Lemma 3.4.5), we have for $d > k_0''$

$$\mathbb{E}\left[u_\pi(t)\right] \leq \begin{cases} \bar{C}_1 L_1(\delta)\sqrt{1+d(t)}\, e^{-\bar{C}_0 t^{1/(4+2r^*)}} + \bar{C}_1 L_1(\delta)\, e^{-\bar{C}_0 \frac{t^{1/(2+r^*)}}{\log t}}, & k_0'' \leq d(t) < d, \\ \bar{C}_1 L_1(\delta)\sqrt{1+d}\, e^{-\bar{C}_0 \frac{t}{d^{3+2r^*}}} + \bar{C}_1 L_1(\delta) e^{-\bar{C}_0 \frac{t}{d^{2(1+r^*)}\log d}}, & d(t) = d. \end{cases} \tag{3.4.66}$$

If instead Assumption 3.2.2 holds, retaining $k_0'', \bar{C}_1, \bar{C}_0$ but switching $d(t)$ according to (3.4.65), *we have*

$$\mathbb{E}\left[u_\pi(t)\right] \leq \begin{cases} \bar{C}_1 L_2(\delta)\sqrt{1+d(t)}\, e^{-\bar{C}_0 t^{1/(1+2r^*)}} + \bar{C}_1 L_2(\delta)\, e^{-\bar{C}_0 \frac{t^{1/(1+2r^*)}}{\log t}}, & k_0'' \leq d(t) < d, \\ \bar{C}_1 L_2(\delta)\sqrt{1+d}\, e^{-\bar{C}_0 \frac{t}{d^{2r^*}}} + \bar{C}_1 L_2(\delta) e^{-\bar{C}_0 \frac{t}{d^{2r^*}\log d}}, & d(t) = d. \end{cases} \tag{3.4.67}$$

Proof. The proof technique is similar to that of Lemma 3.4.7, so we merely sketch the common parts of the argument.

Suppose Assumption 3.2.1 holds, and recall (3.4.28) holds for arbitrary $d' \in \{k_0'', \dots d\}$. Set $d' = d(t)$, with $d(t)$ as in (3.4.65), and consider first the case $d' < d$. Setting $A = 2d'\frac{\overline{\sigma}M}{\underline{\sigma}} \geq A_0$ (by choice of k_0''), we apply (3.4.28) exactly as in (3.4.29) to show,

$$\begin{aligned} -t\frac{\Lambda^{(d')}}{A} &= -t\frac{1}{2d'\frac{\overline{\sigma}M}{\underline{\sigma}}(a^{(d')})^2} \leq -t\frac{b_0^2\underline{\sigma}}{2(d')^{1+2(1+r^*)}\left(\overline{\sigma}M\right)^3} \leq -C_0' t^{1/(4+2r^*)}, \\ -t\frac{\delta'}{T^{(d')}} &\leq -t\frac{\delta'}{1+(d')^{2(1+r^*)}\left(\frac{\overline{\sigma}M}{b_0}\right)^2\log\left(2d'\right)} \leq -C_0'\frac{t^{1/(2+r^*)}}{\log t}, \end{aligned} \tag{3.4.68}$$

for a constant $C_0' > 0$ that does not depend on d, d', r^*. We note the discrepancy of orders in the first and second line of (3.4.68) comes from the extra d'-dependence in the first term, which was not present in (3.4.29).

Fix $x \in \mathbb{R}_+^d$. Arguments preceding (3.4.31) remain valid here: Apply Lemma 3.4.6 with $A = 2d' \frac{\overline{\sigma} M}{\underline{\sigma}}$, using (3.4.68) instead of (3.4.29), to obtain

$$\mathbb{E}\left[u(x,t)\right] \le u(x,0)\left[1+\exp\left(\frac{\|x|_{\mathrm{d}(t)}\|_\infty}{2d(t)\overline{\sigma}Ma^{(\mathrm{d}(t))}}\right)\right]\, e^{-C''C_0' t^{1/(4+2r^*)}} + u(x,0)e^{-C''C_0'\frac{t^{1/(2+r^*)}}{\log t}}$$
$$+\, u(x,0)\,\lambda^{\frac{C_0'}{4}\frac{t^{1/(2+r^*)}}{\log t}-1} + C'\|x\|_{\infty,\delta}\,(\beta/\delta)^{t^{1/(4+2r^*)}}, \qquad \text{for} \quad k_0'' \le \mathrm{d}(t) < d. \quad (3.4.69)$$

As in the proof of (3.4.31), $d(t) \ge \left(\frac{8(4+2r^*)}{C_0' e}\right)^2$ implies $t \ge \frac{4T^{(d(t))}}{\delta}$.

Applying (3.4.69) conditional on $x = X(\infty)$, taking expectations and applying Lemma 3.4.8 to bound the expectations of associated functionals of $X(\infty)$ produces

$$\begin{aligned}
&\mathbb{E}\left[u_\pi(t)\right] = \mathbb{E}\left[u(X(\infty),t)\right]\\
&\le \mathbb{E}\left[u(X(\infty),0)\left[1+\exp\left(\frac{\|X|_{\mathrm{d}(t)}(\infty)\|_\infty}{2d(t)\overline{\sigma}Ma^{(\mathrm{d}(t))}}\right)\right]\right]\, e^{-C''C_0' t^{1/(4+2r^*)}}\\
&+ \mathbb{E}\left[u(X(\infty),0)\right] e^{-C''C_0'\frac{t^{1/(2+r^*)}}{\log t}}\\
&\qquad + \mathbb{E}\left[u(X(\infty),0)\right]\,\lambda^{\frac{C_0'}{4}\frac{t^{1/(2+r^*)}}{\log t}-1} + C'\,\mathbb{E}\left[\|X(\infty)\|_{\infty,\delta}\right]\,(\beta/\delta)^{t^{1/(4+2r^*)}}\\
&\le \mathbb{E}\left[u(X(\infty),0)\left[1+\exp\left(\frac{\|X|_{\mathrm{d}(t)}(\infty)\|_\infty}{2d(t)\overline{\sigma}Ma^{(\mathrm{d}(t))}}\right)\right]\right]\, e^{-C''C_0' t^{1/(4+2r^*)}}\\
&\qquad + C'''L_1(\delta)\left(e^{-C''C_0'\frac{t^{1/(2+r^*)}}{\log t}} + \lambda^{\frac{C_0'}{4}\frac{t^{1/(2+r^*)}}{\log t}-1} + C'\,(\beta/\delta)^{t^{1/(4+2r^*)}}\right)\\
&\le 2C'''L_1(\delta)\sqrt{1+d(t)}\, e^{-C''C_0' t^{1/(4+2r^*)}}\\
&\qquad + C'''L_1(\delta)\left(e^{-C''C_0'\frac{t^{1/(2+r^*)}}{\log t}} + \lambda^{\frac{C_0'}{4}\frac{t^{1/(2+r^*)}}{\log t}-1} + C'\,(\beta/\delta)^{t^{1/(4+2r^*)}}\right),
\end{aligned} \qquad (3.4.70)$$

for $k_0'' \le \mathrm{d}(t) < d$, where $L_1(\delta)$ is defined in Lemma 3.4.8. The second inequality above follows from (3.4.42) and (3.4.43). In the final line we used the Cauchy-Schwarz inequality, the observation that $(1+e^z)^2 \le 4e^{2z}$ for $z \ge 0$, and (3.4.41) and (3.4.43). This proves the first line in (3.4.66), with

$$\bar{C}_1 = C'''\left((2+C') \vee (1+\lambda^{-1})\right) \qquad (3.4.71)$$

and

$$\bar{C}_0 = C''C_0' \wedge \frac{C_0'}{4}\log\frac{1}{\lambda} \wedge \log\frac{\delta}{\beta}. \qquad (3.4.72)$$

We now consider $d(t) = d$, which implies $t \geq d^{4+2r^*}$. Setting $A = 2d\frac{\overline{\sigma}M}{\underline{\sigma}} \geq A_0$, once again we use (3.4.28) with $d' = d$ to show,

$$
\begin{aligned}
-t\frac{\Lambda^{(d')}}{A} &= -t\frac{1}{2d\frac{\overline{\sigma}M}{\underline{\sigma}}(a^{(d)})^2} \leq -t\frac{b_0^2\underline{\sigma}}{2d^{1+2(1+r^*)}\left(\overline{\sigma}M\right)^3} \leq -C_0'\frac{t}{d^{1+2(1+r^*)}}, \\
-t\frac{\delta'}{T^{(d)}} &\leq -t\frac{\delta'}{1+d^{2(1+r^*)}\left(\frac{\overline{\sigma}M}{b_0}\right)^2\log(2d)} \leq -C_0'\frac{t}{d^{2(1+r^*)}\log d}.
\end{aligned} \tag{3.4.73}
$$

The second line of (3.4.66) now follows using Lemma 3.4.6 and Lemma 3.4.8 with $A = 2d\frac{\overline{\sigma}M}{\underline{\sigma}}$ via calculations exactly like (3.4.70), using (3.4.73) instead of (3.4.68).

To prove (3.4.67), i.e. supposing Assumption 3.2.2 holds, we simply use Lemma 3.4.7: Set $x = X(\infty)$ and $A = \max\left\{2\frac{\overline{\sigma}M}{\underline{\sigma}}, A_0\right\}$ in (3.4.27) then take expectations with respect to $X(\infty)$. Result (3.4.67) now follows in a manner perfectly analogous to (3.4.70), using (3.4.44), (3.4.45), (3.4.46) instead of (3.4.41), (3.4.42), (3.4.43). $\square$

With Lemma 3.4.7 and Lemma 3.4.9 in hand, we are now ready to prove Theorem 3.2.1 via (3.4.40).

Proof of Theorem 3.2.1. Fix any $\beta \in (\alpha, 1)$ and $\delta \in (\beta, 1)$. Fix $B \in (0, \infty)$ and fix any $x \in \mathcal{S}(b, B)$. First we consider the case in which Assumption 3.2.1 holds. Since $d(t)$ of (3.4.66) differs slightly from $\ell(t)$ of Lemma 3.4.7, we must take a little care to match the convergence rates appropriately.

Recall $\ell(t)$ of Lemma 3.4.7 is given as $\ell(t) = d \wedge \lfloor t^{1/(3+2r^*)} \rfloor$. Recall from the statement of that lemma the term $k_0' = k_0'(r^*) = k_0 \vee \left(\frac{8(3+2r^*)}{C_0'e}\right)^2$. Then $\left(k_0 \vee \left(\frac{8(3+2r^*)}{C_0'e}\right)^2 + 1\right)^{3+2r^*} \leq t < d^{3+2r^*}$ implies $k_0' \leq \ell(t) < d$. As a result we have directly from the first line of (3.4.26), using $A = A_0$ (A_0 defined in Lemma 3.4.5)

$$
\begin{aligned}
\mathbb{E}\left[u(x,t)\right] &\leq C_1\left(u(x,0)e^{\frac{\|x|_{\ell(t)}\|_\infty}{A_0\underline{\sigma}a^{(\ell(t))}}} + \|x\|_{\infty,\delta}\right) e^{-\frac{C_0}{A_0}t^{1/(3+2r^*)}} + C_1 u(x,0)\, e^{-C_0\frac{t^{1/(3+2r^*)}}{\log t}} \\
&\leq C_1'\|x\|_\infty\left(e^{\frac{\|x|_{\ell(t)}\|_\infty}{A_0\underline{\sigma}a^{(\ell(t))}}} + 1\right) e^{-\frac{C_0}{A_0}t^{1/(3+2r^*)}} + C_1'\|x\|_\infty\, e^{-C_0\frac{t^{1/(3+2r^*)}}{\log t}} \\
&\leq C_1'\|x\|_\infty\left(e^{B/\underline{\sigma}^2} + 1\right) e^{-\frac{C_0}{A_0}t^{1/(3+2r^*)}} + C_1'\|x\|_\infty\, e^{-C_0\frac{t^{1/(3+2r^*)}}{\log t}},
\end{aligned} \tag{3.4.74}
$$

for $\left(k_0 \vee \left(\frac{8(3+2r^*)}{C_0' e}\right)^2 + 1\right)^{3+2r^*} \le t < d^{3+2r^*}$, where C_1' is a constant not depending on d, x. The second-last line above follows from the observation $u(x,0) \le \frac{\beta}{1-\beta}\|x\|_\infty$. In the final line we have used $A_0 \ge 1$, $x \in \mathcal{S}(b,B)$ and $\sum_{j=1}^{\ell(t)} \left(R|_{\ell(t)}\right)^{-1}_{ij} \ge 1$ for $1 \le i \le \ell(t)$ to get from the definition of $a^{(k)}$,

$$\frac{\|x|_{\ell(t)}\|_\infty}{A_0 \underline{\sigma} a^{(\ell(t))}} \le \underline{\sigma}^{-2}\, \underline{b}^{(\ell(t))} \|x|_{\ell(t)}\|_\infty \le \underline{\sigma}^{-2} B.$$

Now recall $d(t) = d \wedge \lfloor t^{1/(4+2r^*)} \rfloor$, which by definition gives $d(t) \le \ell(t)$. Therefore, recalling k_0'' from Lemma 3.4.9, $\left(k_0''^2 + 1\right)^{4+2r^*} \le t < d^{3+2r^*}$ implies $k_0 \vee \left(\frac{8(4+2r^*)}{C_0' e}\right)^2 \le d(t) \le \ell(t) < d$. Hence, we combine (3.4.40), (3.4.74) with the first line of (3.4.66) to obtain

$$\begin{aligned}
&\mathbb{E}\left[\|\left(X(x,t) - X(X(\infty),t)\right)\|_{1,\beta}\right] \\
&\quad \le \mathbb{E}\left[u_\pi(t)\right] + C_1' \|x\|_\infty \left(e^{B/\underline{\sigma}^2} + 1\right) e^{-\frac{C_0}{A_0} t^{1/(3+2r^*)}} + C_1' \|x\|_\infty\, e^{-C_0 \frac{t^{1/(3+2r^*)}}{\log t}} \\
&\quad \le \bar{C}_1 L_1(\delta) \sqrt{1 + t^{1/(4+2r^*)}}\, e^{-\bar{C}_0 t^{1/(4+2r^*)}} + \bar{C}_1 L_1(\delta)\, e^{-\bar{C}_0 \frac{t^{1/(2+r^*)}}{\log t}} \\
&\qquad + C_1' \|x\|_\infty \left(e^{B/\underline{\sigma}^2} + 1\right) e^{-\frac{C_0}{A_0} t^{1/(3+2r^*)}} + C_1' \|x\|_\infty\, e^{-C_0 \frac{t^{1/(3+2r^*)}}{\log t}},
\end{aligned} \tag{3.4.75}$$

for $\left(k_0''^2 + 1\right)^{4+2r^*} \le t < d^{3+2r^*}$. Now if $d^{3+2r^*} \le t < d^{4+2r^*}$ the bound on $\mathbb{E}\left[u_\pi(t)\right]$ in the first line of (3.4.66) continues to hold, and the bound on $\mathbb{E}\left[u(x,t)\right]$ from the second line of (3.4.26) is now valid. Thus, we have

$$\begin{aligned}
&\mathbb{E}\left[\|\left(X(x,t) - X(X(\infty),t)\right)\|_{1,\beta}\right] \\
&\quad \le \bar{C}_1 L_1(\delta) \sqrt{1 + t^{1/(4+2r^*)}}\, e^{-\bar{C}_0 t^{1/(4+2r^*)}} + \bar{C}_1 L_1(\delta)\, e^{-\bar{C}_0 \frac{t^{1/(2+r^*)}}{\log t}} \\
&\qquad + C_1' \|x\|_\infty \left(e^{B/\underline{\sigma}^2} + 1\right) e^{-\frac{C_0}{A_0} \frac{t}{d^{2(1+r^*)}}} + C_1' \|x\|_\infty\, e^{-C_0 \frac{t}{d^{2(1+r^*)} \log d}} \\
&\quad \le \bar{C}_1 L_1(\delta) \sqrt{1 + t^{1/(4+2r^*)}}\, e^{-\bar{C}_0 t^{1/(4+2r^*)}} + \bar{C}_1 L_1(\delta)\, e^{-\bar{C}_0 \frac{t^{1/(2+r^*)}}{\log t}} \\
&\qquad + C_1' \|x\|_\infty \left(e^{B/\underline{\sigma}^2} + 1\right) e^{-\frac{C_0}{A_0} t^{1/(3+2r^*)}} + C_1' \|x\|_\infty\, e^{-3C_0 \frac{t^{1/(3+2r^*)}}{\log t}},
\end{aligned} \tag{3.4.76}$$

for $d^{3+2r^*} \le t < d^{4+2r^*}$, where the final inequality follows from $\frac{t}{d^{2(1+r^*)}} \ge t^{1/(3+2r^*)}$ and $\log d \le \frac{\log t}{3+2r^*} \le \frac{\log t}{3}$. The first line in (3.2.9) follows from (3.4.75), (3.4.76) by taking $\beta = \sqrt{\alpha}$ and $\delta = \alpha^{1/4}$ after keeping only leading-order terms in the above bounds, for simplicity.

To prove the second line in (3.2.9): Note the second lines of (3.4.66) and (3.4.26) remain valid for all $t \geq d^{4+2r(d)}$. Applying those results to (3.4.40) and otherwise proceeding as in the lead-up to (3.4.76)

$$
\begin{aligned}
&\mathbb{E}\left[\| \left(X(x,t) - X(X(\infty),t)\right) \|_{1,\beta}\right] \\
&\quad \leq C_1 \|x\|_\infty e^{\frac{\|x\|_\infty}{A_0 \underline{\sigma} a(d)}} \, e^{-\frac{C_0}{A_0} \frac{t}{d^{2(1+r^*)}}} + C_1 \|x\|_\infty e^{-C_0 \frac{t}{d^{2(1+r^*)} \log d}} \\
&\qquad + C_1 L_1(\delta) \sqrt{1+d} \, e^{-\bar{C}_0 \frac{t}{d^{3+2r^*}}} + C_1 L_1(\delta) e^{-C_0 \frac{t}{d^{2(1+r^*)} \log d}} \\
&\quad \leq C_1 \|x\|_\infty e^{B/\underline{\sigma}^2} \, e^{-\frac{C_0}{A_0} \frac{t}{d^{2(1+r^*)}}} + C_1 \|x\|_\infty e^{-C_0 \frac{t}{d^{2(1+r^*)} \log d}} \\
&\qquad + C_1 L_1(\delta) \sqrt{1 + t^{1/(4+2r^*)}} \, e^{-\bar{C}_0 \frac{t}{d^{3+2r^*}}} + C_1 L_1(\delta) e^{-C_0 \frac{t}{d^{2(1+r^*)} \log d}},
\end{aligned}
\tag{3.4.77}
$$

for $t \geq d^{4+2r^*}$ and constants $C_0, C_1 > 0$ not depending on d, r^* or B. The second line in (3.2.9) follows from (3.4.77) by taking $\beta = \sqrt{\alpha}$ and $\delta = \alpha^{1/4}$ and by keeping only leading-order terms in (3.4.77).

(3.2.10) follows in identical fashion, using (3.4.27) instead of (3.4.26) and (3.4.67) instead of (3.4.66). We therefore omit the proof. $\square$

3.5 Proofs: Perturbations from stationarity for the Symmetric Atlas Model

Proof of Theorem 3.3.6. The almost sure existence and representation (3.3.7) of the derivative is a consequence of Andres (2009, Theorem 1.2), as we now show. The cited theorem proves the almost sure existence of the derivative process up to the first time X hits a corner of the orthant $\mathbb{R}^d_+$. Since the Atlas model does not hit corners by Sarantsev (2015, Theorem 1.9), the derivative $\eta^{i_0}(x,t)$ exists almost surely for any $t \in [0, \infty)$.

For $1 \leq i \leq d$, the vector v_i of Andres (2009, Theorem 1.2) is the ith column of R here, denoted $R^{(i)}$, and the ith inward normal n_i of $\mathbb{R}^d_+$ is the standard basis vector e_i. Terms $\frac{\partial}{\partial x_j} b(X(x,t))$ of Andres (2009, Theorem 1.2) are all zero here, since the drift $b(X(x,t)) = \mu t$ does not depend on x. For $1 \leq i \leq d$, define vectors $\left(R^{(i)}\right)^\perp$ and $e_i^\perp$, orthogonal to $R^{(i)}$ and e_i respectively, by equation (1.1) of Andres (2009) such that these vectors lie in $\operatorname{span}\{R^{(i)}, e_i\}$. For $d \geq 3$, extend $e_i, e_i^\perp$ by the vectors $\{n_i^j\}_{3 \leq j \leq d}$ to an orthonormal basis of $\mathbb{R}^d_+$.

From Andres (2009, Theorem 1.2), defining $S_k^{i_0}(x) = \eta^{i_0}(x, \tau_k)$ for $k \geq 0$,

$$S_{k+1}^{i_0} = \left\langle S_k^{i_0}(x), \left(R^{(i_{k+1})}\right)^{\perp} \right\rangle e_{i_{k+1}}^{\perp} + \sum_{j=3}^{d} \left\langle S_k^{i_0}(x), n_{i_{k+1}}^j \right\rangle n_{i_{k+1}}^j, \tag{3.5.1}$$

and $\eta^{i_0}(x,t)$ is constant on $t \in [\tau_k, \tau_{k+1})$. Moreover,

$$S_k^{i_0} = \left\langle S_k^{i_0}(x), e_{i_{k+1}} \right\rangle e_{i_{k+1}} + \left\langle S_k^{i_0}(x), e_{i_{k+1}}^{\perp} \right\rangle e_{i_{k+1}}^{\perp} + \sum_{j=3}^{d} \left\langle S_k^{i_0}(x), n_{i_{k+1}}^j \right\rangle n_{i_{k+1}}^j. \tag{3.5.2}$$

In the above representations, the sum $\sum_{j=3}^{d}$ is taken to be zero if $d = 1, 2$. From (3.5.1), (3.5.2) and Andres (2009, Lemma 1.7),

$$\begin{aligned} S_{k+1}^{i_0} - S_k^{i_0} &= \left\langle S_k^{i_0}(x), \left(R^{(i_{k+1})}\right)^{\perp} \right\rangle e_{i_{k+1}}^{\perp} - \left\langle S_k^{i_0}(x), e_{i_{k+1}} \right\rangle e_{i_{k+1}} - \left\langle S_k^{i_0}(x), e_{i_{k+1}}^{\perp} \right\rangle e_{i_{k+1}}^{\perp} \\ &= - \left\langle S_k^{i_0}(x), e_{i_{k+1}} \right\rangle R^{(i_{k+1})} \end{aligned} \tag{3.5.3}$$

which proves (3.3.7).

It remains only to prove the random walk representation (3.3.8). Define the $\mathbb{R}_+^{d+2}$-valued functions $u(\cdot)$ and $v(\cdot)$ as follows: $v_j(t) := \mathbb{P}_{\Theta(x), i_0}(W(t) = j)$, $j \in \{0, \ldots, d+1\}$. Set $u_j(t) := \eta_j^{i_0}(x,t)$ for $j = 1, \ldots, d$ and define $u_0(\cdot), u_{d+1}(\cdot)$ iteratively by $u_0(\tau_{k+1}) = u_0(\tau_k) + \frac{1}{2} u_1(\tau_k) \mathbb{1}_{i_{k+1}=1}$, $u_{d+1}(\tau_{k+1}) = u_{d+1}(\tau_k) + \frac{1}{2} u_d(\tau_k) \mathbb{1}_{i_{k+1}=d}$, with $u_0(\cdot), u_{d+1}(\cdot)$ constant on $t \in [\tau_k, \tau_{k+1})$ for $k \geq 0$.

Using (3.3.7) for $u(\cdot)$ and the defining properties of $RW(\Theta(x), i_0)$ for $v(\cdot)$, note that for any $k \geq 0$, $u(t) = u(\tau_k)$ and $v(t) = v(\tau_k)$ for all $t \in [\tau_k, \tau_{k+1})$. Hence, we only need to show that $u(\tau_k) = v(\tau_k), k \geq 0$. This follows from the fact that both $\{u(\tau_k)\}_{k \geq 0}$ and $\{v(\tau_k)\}_{k \geq 0}$ are solutions to the recursive equation in $\{w(k)\}_{k \geq 0}$: $w(0) = e_{i_0}$ and for $k \geq 0$, with the fixed integer sequence $\{i_k\}_{k \geq 0}$,

$$\begin{aligned} w_j(k+1) = &\left(w_j(k) + \frac{1}{2} w_{j-1}(k) \right) \mathbb{1}_{i_{k+1}=j-1} \\ &+ \left(w_j(k) + \frac{1}{2} w_{j+1}(k) \right) \mathbb{1}_{i_{k+1}=j+1} + w_j(k) \mathbb{1}_{i_{k+1} \neq j, j \pm 1}, \quad 1 \leq j \leq d, \end{aligned} \tag{3.5.4}$$

and

$$w_0(k+1) = w_0(k) + \frac{1}{2}w_1(k)\mathbb{1}_{i_{k+1}=1}, \qquad w_{d+1}(k+1) = w_{d+1}(k) + \frac{1}{2}w_d(k)\mathbb{1}_{i_{k+1}=d}. \tag{3.5.5}$$

Note an inductive argument implies $\sum_{j=0}^{d+1} w_j(k) = 1$ for all $k \geq 0$ for any solution to (3.5.4), (3.5.5) with $w(0) = e_{i_0}$.

(3.5.4), (3.5.5) hold for $\{u(\tau_k)\}_{k\geq 0}$ by (3.3.7) and for $\{v(\tau_k)\}_{k\geq 0}$ by the definition of $RW(\Theta(x), i_0)$. Since $u(0) - v(0) = 0$, the sequence $\{i_k\}_{k\geq 0}$ is common to u and v, and (3.5.4), (3.5.5) are linear recursive equations in $w(\cdot)$, we have $u(\tau_k) - v(\tau_k) = 0$ for all $k \geq 0$. This proves (3.3.8). □

Proof of Theorem 3.3.2. The proof consists of analyzing two regimes: $t < d^{16/3}$ and $t > t_0'' d^4 \log(2d)$. In the former regime, we show that the probability of any of the first $m(t)$ coordinates of X not hitting zero sufficiently often is well-controlled by Lemma 3.4.5, for appropriately chosen time-dependent integer $m(t)$. On the other hand, if each of the first $m(t)$ coordinates of X makes a large number of visits to zero, then the random walk W in the derivative representation of Theorem 3.3.6 makes a large number of jumps, and consequently, has a higher chance of getting absorbed in 0 or $d+1$ by time t. In this case, we bound the right hand side of Corollary 3.3.7 using the probability that a simple random walk does not hit 0 within a certain number of steps. For $t > t_0'' d^4 \log(2d)$, we use the approach of Banerjee and Budhiraja (2020) via contractions in L^1 distance between the synchronously coupled RBMs.

Note that the Atlas model X satisfies $b^{(d')} = -\left(R|_{d'}\right)^{-1} \mu|_{d'} = \left\{(R|_{d'})^{-1}_{i1}\right\}_{i=1}^{d'} > 0$ for every $d' \in \{1, \ldots, d\}$, and IV of Assumption 3.2.1 holds with $k_0 = 1$ and $\underline{\sigma} = \overline{\sigma} = \sqrt{2}$. Therefore we may apply Lemma 3.4.5, and in preparation we first calculate the quantities $a^{(d')}, T^{(d')}, \Lambda^{(d')}$ for $d' \in \{1, \ldots, d\}$.

Recalling that $b^{(d')}$ is the first column of $(R|_{d'})^{-1}$ and computing the row sums of $(R|_{d'})^{-1}$ from (3.3.2) with d' in place of d gives

$$a^{(d')} = \max_{1\leq i\leq d'} \frac{1}{b_i^{(d')}} \sum_{j=1}^{d'} (R|_{d'})^{-1}_{ij} \sigma_j = \max_{1\leq i\leq d'} \frac{\sqrt{2}i(d'+1-i)}{2\left(1 - \frac{i}{d'+1}\right)} = \frac{d'(d'+1)}{\sqrt{2}}. \tag{3.5.6}$$

Plugging this into the definitions of $T^{(d')}, \Lambda^{(d')}$ in (3.4.18) and applying Lemma 3.4.5, we obtain $A_0 \geq 1$ not depending on d, d' such that for any $d' \in \{1, \ldots, d\}$, $t \geq 4\left(1 + \frac{1}{2}(d'(d'+1))^2 \log(2d')\right)/\delta'$ and $A \geq A_0$,

$$\begin{aligned}
&\mathbb{P}\left[\mathcal{N}_{d'}(x,t) < t\frac{\delta'}{4\left(1 + \frac{1}{2}(d'(d'+1))^2 \log(2d')\right)}\right] \\
&\leq \exp\left(-t\frac{\delta' C''}{1 + \frac{1}{2}(d'(d'+1))^2 \log(2d')}\right) + \exp\left(-t\frac{2C''}{A(d'(d'+1))^2}\right)\left\{1 + \exp\left(\frac{\|x|_{d'}\|_\infty}{Ad'(d'+1)}\right)\right\},
\end{aligned} \tag{3.5.7}$$

where $\delta', C'' > 0$ and $A_0 \geq 1$ do not depend on d', d. We now consider $d' = m(t)$, where $m(t) \in \{1, \ldots, d\}$ will be a time-dependent integer to be determined later. Recall $\tau_0^* := \inf\{s \geq 0 : W(s) = 0\}$. For any integer $n(t)$ such that $1 \leq n(t) < m(t)$ for t large enough that (3.5.7) holds (a time t_0 to be determined below) and with $N(t) = t\frac{\delta'}{4\left(1+\frac{1}{2}(d'(d'+1))^2 \log(2d')\right)}$ we have for $i \in \{1, \ldots, n(t)\}$,

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{P}_{\Theta(x),i}\left(\tau_0^* > t,\ \max_{0 \leq s \leq t} W(s) < m(t)\right)\right] \\
&\leq \mathbb{E}\left[\mathbb{P}_{\Theta(x),i}\left(\tau_0^* > t,\ \max_{0 \leq s \leq t} W(s) < m(t)\right)\mathbb{1}_{\mathcal{N}_{m(t)}(x,t) \geq N(t)}\right] + \mathbb{P}\left[\mathcal{N}_{m(t)}(x,t) < N(t)\right] \\
&\leq 12\frac{n(t)}{\sqrt{N(t)}} + \mathbb{P}\left[\mathcal{N}_{m(t)}(x,t) < N(t)\right] \\
&\leq 12\frac{n(t)}{\sqrt{N(t)}} + \exp\left(-4C'' N(t)\right) \\
&\qquad + \exp\left(-8C'' N(t)\frac{1 + \frac{1}{2}(m(t)(m(t)+1))^2 \log(2m(t))}{\delta' A (m(t)(m(t)+1))^2}\right)\left\{1 + \exp\left(\frac{\|x|_{m(t)}\|_\infty}{Am(t)(m(t)+1)}\right)\right\} \\
&\leq 12\frac{n(t)}{\sqrt{N(t)}} + \exp\left(-4C'' N(t)\right) + \exp\left(-\frac{4C''}{\delta' A}N(t)\right)\left\{1 + \exp\left(\frac{\|x|_{m(t)}\|_\infty}{Am(t)(m(t)+1)}\right)\right\}.
\end{aligned} \tag{3.5.8}$$

The second inequality above follows from (3.3.9) with $m = m(t)$ and a standard bound on the probability that a simple random walk started from $i \in \{1, \ldots, n(t)\}$ has not hit 0 after $N(t)$ steps (e.g. Levin, D. and Peres, Y. and Wilmer, E. (2017) Theorem 2.17). The third inequality applies (3.5.7) with $d' = m(t)$ and $t = N(t)(\delta')^{-1} 4\left(1 + \frac{1}{2}(m(t)(m(t)+1))^2 \log(2m(t))\right)$.

Now for $i \in \{1, \dots, n(t)\}$ such that $W(0) = i$, the event $\{\tau_0^* > t,\ \max_{0 \le s \le t} W(s) \ge m(t)\}$ implies the walk W has taken at least $m(t) - n(t)$ steps without hitting 0 or $d+1$, where it is absorbed. Thus for all $i \in \{1, \dots, n(t)\}$,

$$\mathbb{P}_{\Theta(x),i}\left(\tau_0^* > t,\ \max_{0 \le s \le t} W(s) \ge m(t)\right) \le 12 \frac{n(t)}{\sqrt{m(t) - n(t)}}. \tag{3.5.9}$$

We now set $m(t)$ so that the bounds in (3.5.8), (3.5.9) are of the same order. Fix $\epsilon \in (0, 1/4)$ to be chosen later. Set $m(t) = d \wedge \lfloor t^{1/4-\epsilon} \rfloor$. There exists a $t_0(\epsilon) > 0$ not depending on d such that

$$N(t) = \frac{t\delta'}{1 + \frac{1}{2}\left(m(t)(m(t)+1)\right)^2 \log(2m(t))} \ge \frac{t\delta'}{t^{1-4\epsilon} \log(2t^{1/4-\epsilon})} \ge t^{3\epsilon}, \tag{3.5.10}$$

for $t \ge t_0(\epsilon)$. From this, we conclude that if t is chosen such that $d \ge \lfloor t^{1/4-\epsilon} \rfloor \ge \lfloor t_0(\epsilon)^{1/4-\epsilon} \rfloor$ and $n(t) \le m(t)/2$, the dominating term in (3.5.8) is of order $n(t)t^{-\frac{3}{2}\epsilon}$ and the dominating term in (3.5.9) is of order $n(t)t^{-\frac{1}{8}+\frac{\epsilon}{2}}$.

Setting $\epsilon = \frac{1}{16}$ matches these orders, at $n(t)t^{-\frac{3}{32}}$. Therefore we set $t_0 = t_0(1/16)$ and define

$$m(t) = d \wedge \lfloor t^{1/4-\epsilon} \rfloor = d \wedge \lfloor t^{3/16} \rfloor. \tag{3.5.11}$$

We are now ready to prove (3.3.5). Choose and fix any $n(\cdot)$ as in the statement of the theorem, and recall the definition of $t_0^{(n)}$ given there. We have by (3.5.8), (3.5.9), (3.5.10) for any $1 \le i \le n(t)$ and $t \ge t_0^{(n)}$, which implies $2n(t) \vee \lfloor t_0^{3/16} \rfloor \le m(t) \le d$,

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{P}_{\Theta(x),i}\left(\tau_0^* > t\right)\right] \\
&= \mathbb{E}\left[\mathbb{P}_{\Theta(x),i}\left(\tau_0^* > t,\ \max_{0 \le s \le t} W(s) \ge m(t)\right)\right] + \mathbb{E}\left[\mathbb{P}_{\Theta(x),i}\left(\tau_0^* > t,\ \max_{0 \le s \le t} W(s) < m(t)\right)\right] \\
&\le 12 \frac{n(t)}{\sqrt{m(t) - n(t)}} + 12 \frac{n(t)}{\sqrt{N(t)}} + \exp\left(-4C'' N(t)\right) \\
&\qquad + \exp\left(-\frac{4C''}{\delta' A} N(t)\right)\left\{1 + \exp\left(\frac{\|x|_{m(t)}\|_\infty}{Am(t)(m(t)+1)}\right)\right\} \\
&\le 12(1+\sqrt{2}) \frac{n(t)}{\sqrt{m(t)}} + \exp\left(-4C'' N(t)\right) + \exp\left(-\frac{4C''}{\delta' A} N(t)\right)\left\{1 + \exp\left(\frac{\|x|_{m(t)}\|_\infty}{Am(t)(m(t)+1)}\right)\right\}.
\end{aligned} \tag{3.5.12}$$

This holds for any $A \geq A_0$ given in (3.5.7). In the final inequality we used $m(t) \geq 2n(t)$ implies $\sqrt{m(t)-n(t)} \geq 2^{-1/2}\sqrt{m(t)}$, and $N(t) \geq t^{3/16} \geq m(t)$ by (3.5.11) with the chosen $\epsilon = 1/16$.

For $x, \tilde{x} \in \mathbb{R}_+^d$ with $x > 0$ and $t \geq 0$, by Corollary 3.3.7, with $\gamma(u) = x + u(\tilde{x} - x), u \in [0,1]$,

$$\|X(\tilde{x},t) - X(x,t)\|_1 \leq \sum_{i=1}^{n(t)} |(\tilde{x}-x)_i| \int_{[0,1)} \mathbb{P}_{\Theta(\gamma(u)),i}(\tau_0^* > t)\, du + \sum_{i=n(t)+1}^{d} |(\tilde{x}-x)_i|. \quad (3.5.13)$$

Applying (3.5.12) to (3.5.13) and using $N(t) \geq m(t)$, we have for $2n(t) \vee \lfloor t_0^{3/16} \rfloor \leq m(t) \leq d$,

$$\begin{aligned}
&\mathbb{E}\left[\|X(\tilde{x},t) - X(x,t)\|_1\right] \\
&\leq \left[12(1+\sqrt{2})\|x-\tilde{x}\|_1\right] \frac{n(t)}{\sqrt{m(t)}} + \|x-\tilde{x}\|_1 \exp\left(-4C'' m(t)\right) \\
&+ \left[\|x-\tilde{x}\|_1 \int_{[0,1)} \left\{1 + \exp\left(\frac{\|\gamma(u)|_{m(t)}\|_\infty}{Am(t)(m(t)+1)}\right)\right\} du\right] \exp\left(-\frac{4C''}{\delta' A} m(t)\right) \\
&+ \sum_{i=n(t)+1}^{d} |(\tilde{x}-x)_i|. \qquad (3.5.14)
\end{aligned}$$

Fix any $Y \in \mathcal{P}(P_1, P_2, \delta)$. Recall $X^Y(\infty) := (X(\infty) + Y|_d)_+$ and $\alpha^Y(\cdot)$ from (3.3.4). Using (3.5.14) conditioned on $x = X(\infty), \tilde{x} = X^Y(\infty)$, then taking expectations, and using the fact

$\|X(\infty) - X^Y(\infty)\|_1 \le \sum_1^d |Y_i| \le \sum_1^\infty |Y_i| = \|Y\|_1$, we have for $2n(t) \vee \lfloor t_0^{3/16} \rfloor \le m(t) \le d$,

$$\begin{aligned}
&\mathbb{E}\left[\|X(X^Y(\infty),t) - X(X(\infty),t)\|_1\right] - \alpha^Y(n(t)) \\
&\le \left[12(1+\sqrt{2})\,\mathbb{E}\left[\|Y\|_1\right]\right] \frac{n(t)}{\sqrt{m(t)}} + \mathbb{E}\left[\|Y\|_1\right] \exp\left(-4C''m(t)\right) \\
&\quad + \mathbb{E}\left[\|Y\|_1 \int_{[0,1)} \left\{1 + \exp\left(\frac{\|\gamma(u)|_{m(t)}\|_\infty}{Am(t)(m(t)+1)}\right)\right\} du\right] \exp\left(-\frac{4C''}{\delta' A} m(t)\right) \\
&\le \left[12(1+\sqrt{2})\,\mathbb{E}\left[\|Y\|_1\right]\right] \frac{n(t)}{\sqrt{m(t)}} + \mathbb{E}\left[\|Y\|_1\right] \exp\left(-4C''m(t)\right) \\
&\quad + \sqrt{\mathbb{E}\left[\|Y\|_1^2\right]} \left[1 + \sqrt{\mathbb{E}\left[\exp\left(\frac{2\|Y|_{m(t)}(\infty)\|_\infty}{Am(t)(m(t)+1)}\right) \exp\left(\frac{2\|X|_{m(t)}(\infty)\|_\infty}{Am(t)(m(t)+1)}\right)\right]}\right] \\
&\qquad\qquad \times \exp\left(-\frac{4C''}{\delta' A} m(t)\right) \\
&\le \left[12(1+\sqrt{2})\,\mathbb{E}\left[\|Y\|_1\right]\right] \frac{n(t)}{\sqrt{m(t)}} + \mathbb{E}\left[\|Y\|_1\right] \exp\left(-4C''m(t)\right) \\
&\quad + \sqrt{\mathbb{E}\left[\|Y\|_1^2\right]} \left[1 + \left(\mathbb{E}\left[\exp\left(\frac{4\|Y|_{m(t)}(\infty)\|_\infty}{Am(t)(m(t)+1)}\right)\right] \mathbb{E}\left[\exp\left(\frac{4\|X|_{m(t)}(\infty)\|_\infty}{Am(t)(m(t)+1)}\right)\right]\right)^{1/4}\right] \\
&\qquad\qquad \times \exp\left(-\frac{4C''}{\delta' A} m(t)\right). \qquad (3.5.15)
\end{aligned}$$

In the second inequality, we used the Cauchy-Schwarz inequality and the observation that for any $m \in \{1, \ldots, d\}$, $\|\gamma(u)|_m\|_\infty = \max_{1 \le i \le m} \left|X_i(\infty) + u(X_i^Y(\infty) - X_i(\infty))\right| \le \|Y|_m(\infty)\|_\infty + \|X|_m(\infty)\|_\infty$ for $u \in [0,1]$.

As $Y \in \mathcal{P}(P_1, P_2, \delta)$, taking $A = 4\max\{A_0, 4\delta^{-1}\}$, where A_0 is given in (3.5.7),

$$\begin{aligned}
\mathbb{E}\left[\|Y\|_1\right] &\le \sqrt{\mathbb{E}\left[\|Y\|_1^2\right]} \le \sqrt{P_1}, \\
\mathbb{E}\left[\exp\left(\frac{4\|Y|_{m(t)}(\infty)\|_\infty}{Am(t)(m(t)+1)}\right)\right] &\le P_2. \qquad (3.5.16)
\end{aligned}$$

Moreover, for the same choice of A, we obtain along the same lines as (3.4.52) using the explicit product form distribution of $X|_{m(t)}(\infty)$ (see (3.3.3)),

$$\begin{aligned}
\mathbb{E}\left[\exp\left(\frac{4\|X|_{m(t)}(\infty)\|_\infty}{Am(t)(m(t)+1)}\right)\right] &\le \mathbb{E}\left[\exp\left(\frac{\|X|_{m(t)}(\infty)\|_\infty}{A_0 m(t)(m(t)+1)}\right)\right] \\
&\le 1 + \frac{m(t)}{A_0 m(t)(m(t)+1) - 1} \le 2. \qquad (3.5.17)
\end{aligned}$$

Note that we cannot refer to Lemma 3.4.8 here since Assumption 3.2.1 does not hold for the Atlas model. Using the above estimates in (3.5.15), we obtain for $2n(t) \vee \lfloor t_0^{3/16} \rfloor \leq m(t) \leq d$,

$$\mathbb{E}\left[\|X(X^Y(\infty),t) - X(X(\infty),t)\|_1\right] \leq \sqrt{P_1}\left[12(1+\sqrt{2})\,\frac{n(t)}{\sqrt{m(t)}} + \exp\left(-4C''m(t)\right)\right]$$
$$+ \sqrt{P_1}\left(1 + (2P_2)^{1/4}\right)\exp\left(-\frac{C''}{\delta' \max\{A_0, 4\delta^{-1}\}}m(t)\right) + \alpha^Y(n(t)). \tag{3.5.18}$$

This proves the first bound in (3.3.5) upon noting that $\frac{C''}{\delta' \max\{A_0, 4\delta^{-1}\}} \geq \frac{C''}{\delta' A_0}\frac{\delta}{\delta+4}$, and for $t_0^{(n)} \leq t < d^{16/3}$ (with $t_0^{(n)}$ as defined in the theorem statement), $2n(t) \vee \lfloor t_0^{3/16} \rfloor \leq m(t) \leq d$.

We now address the case when t is large relative to d by applying results from Banerjee and Budhiraja (2020). Using equation (44) of that reference, plugging in the Standard Atlas model parameter estimates calculated in equation (3.5.6) here (with $d = d'$) and, in the reference, equation (12) and parameters given prior to Theorem 1, we have for any $x, \tilde{x} \in \mathbb{R}_+^d$ with $x > 0$,

$$\mathbb{E}\left[\|X(x,t) - X(\tilde{x},t)\|_1\right] \leq \mathbb{E}\left[\|X(x,t) - X(0,t)\|_1\right] + \mathbb{E}\left[\|X(\tilde{x},t) - X(0,t)\|_1\right]$$
$$\leq C_1\left(\|x\|_1 \exp\left(\frac{C_0'\|x\|_\infty}{A'd^4}\right) + \|\tilde{x}\|_1 \exp\left(\frac{\|C_0'\tilde{x}\|_\infty}{A'd^4}\right)\right)\exp\left(-\frac{C_0}{A'}\frac{t}{d^6\log(2d)}\right), \tag{3.5.19}$$

for all $t \geq t_0'' d^4 \log(2d)$, $A' \geq A_0'$, where $C_0, C_0', C_1, t_0'', A_0' \in (0,\infty)$ are dimension-independent constants. Applying (3.5.19) conditional on $x = X(\infty) > 0$ and $\tilde{x} = X^Y(\infty) \geq 0$ and taking expectations we have

$$\mathbb{E}\left[\|X(X^Y(\infty),t) - X(X(\infty),t)\|_1\right]$$
$$\leq C_1\left(\mathbb{E}\left[\|X(\infty)\|_1 \exp\left(\frac{C_0'\|X(\infty)\|_\infty}{A'd^4}\right)\right] + \mathbb{E}\left[\|X^Y(\infty)\|_1 \exp\left(\frac{\|C_0'X^Y(\infty)\|_\infty}{A'd^4}\right)\right]\right)$$
$$\times \exp\left(-\frac{C_0}{A'}\frac{t}{d^6\log(2d)}\right) \tag{3.5.20}$$

for all $t \geq t_0'' d^4 \log(2d)$, $A' \geq A_0'$. From the explicit distribution of $X(\infty)$ in (3.3.3), for any $A' \geq \max\{A_0', 4C_0'\}$,

$$\mathbb{E}\left[\|X(\infty)\|_1 \exp\left(\frac{C_0'\|X(\infty)\|_\infty}{A'd^4}\right)\right] \leq \sqrt{\mathbb{E}\left[\|X(\infty)\|_1^2\right]}\sqrt{\mathbb{E}\left[\exp\left(\frac{2C_0'\|X(\infty)\|_\infty}{A'd^4}\right)\right]} \leq 2d, \tag{3.5.21}$$

Moreover, as $Y \in \mathcal{P}(P_1, P_2, \delta)$, using $\|X^Y(\infty)\|_1 \leq \|X(\infty)\|_1 + \|Y\|_1$ and $\|X^Y(\infty)\|_\infty \leq \|X(\infty)\|_\infty + \|Y|_d\|_\infty$, we obtain for any $A' \geq \max\{A_0', 2C_0'\delta^{-1}, 4C_0'\}$,

$$\begin{aligned}
&\mathbb{E}\left[\|X^Y(\infty)\|_1 \exp\left(\frac{C_0'\|X^Y(\infty)\|_\infty}{A'd^4}\right)\right] \leq \sqrt{\mathbb{E}\left[\|X^Y(\infty)\|_1^2\right]}\sqrt{\mathbb{E}\left[\exp\left(\frac{2C_0'\|X^Y(\infty)\|_\infty}{A'd^4}\right)\right]}\\
&\leq \sqrt{2\,\mathbb{E}\left[\|X(\infty)\|_1^2\right] + 2\,\mathbb{E}\left[\|Y\|_1^2\right]}\sqrt{\mathbb{E}\left[\exp\left(\frac{2C_0'\|X(\infty)\|_\infty}{A'd^4}\right)\right]\mathbb{E}\left[\exp\left(\frac{2C_0'\|Y|_d\|_\infty}{A'd^4}\right)\right]}\\
&\leq \sqrt{4d^2 + 2P_1}\sqrt{2P_2}.
\end{aligned} \tag{3.5.22}$$

Using (3.5.21) and (3.5.22) in (3.5.20), fixing $A' = \max\{A_0', 2C_0'\delta^{-1}, 4C_0'\}$, we obtain

$$\mathbb{E}\left[\|X(X^Y(\infty), t) - X(X(\infty), t)\|_1\right] \leq 2C_1\sqrt{4d^2 + 2P_1}\sqrt{2P_2}\exp\left(-\frac{C_0}{A'}\frac{t}{d^6\log(2d)}\right), \tag{3.5.23}$$

for $t \geq t_0'' d^4 \log(2d)$, which proves the second bound in (3.3.5), and completes the proof of the theorem. $\square$

CHAPTER 4

Explorations of synchronous-like couplings for approximation

4.1 Introduction

In this chapter we diverge from the previous ones by considering couplings between two stochastic processes for the purpose of using one to approximate the other according to some cost, rather than to evaluate stability questions. We are drawn to do so by questions from surrogate and multifidelity modeling for dynamicals systems in engineering and applied mathematics. There, broadly speaking the goal is to study a complex system by leveraging statistical estimates from a simplified version. Section 4.1.2 will describe briefly this perspective and summarize some work from this author's ongoing collaboration to study laws of rare events using such methods — mostly as a soft introduction to the point of view in this chapter.

Furthermore, we discuss the ways in which these couplings are constructive and computable, in the sense that given time series data for a given stochastic system, the cost-minimizing coupling could be constructed with available statistical methods. However, as computational questions are not the main topic here, we simple recognize this possibility and avoid a more detailed analysis of how to implement such a procedure.

Throughout we make a semi-formal presentation: The section is a series of propositions, remarks and examples that avoid technicalities where possible to focus on the main ideas, which are largely preliminary in nature and the seed of future work.

4.1.1 Definitions and key examples

Throughout the chapter we will examples that are as simple as possible while still illustrating some of the difficulties and results one might expect.

We will work with pairs of random variables (X, Y) taking values in complete, separable metric (Polish) spaces, usually $\mathbb{R}$ or $\mathcal{C}([0,\infty),\mathbb{R})$, which by regular conditioning we can and do assume to be defined on a common probability space Thorisson (2000, Ch 3.4).

Example 4.1.1 (Vanilla diffusion in one dimension)**.** *A diffusion $X \in \mathcal{C}([0,\infty),\mathbb{R})$ is a solution to the Itô stochastic differential equation*

$$dX_t = b(X_t)\,dt + \sigma(X_t)dW_t, \quad t > 0, \tag{4.1.1}$$

where W is a one-dimensional standard Brownian motion. We will suppose Assumptions 4.1.1 hold, which guarantees the solution X is unique for each initial condition $X_0 = x_0 \in \mathbb{R}$ Le Gall (2013, Theorem 8.3, e.g.).

Assumption 4.1.1. *The real-valued functions $b, \sigma \in \mathcal{C}_b^2$, and the diffusion coefficient satisfies the ellipticity condition $|\sigma(x)| \geq C' > 0$ for some C' and all x.*

Next we define the synchronous coupling between two diffusions. This also motivates our informal use of the term 'synchronous-like' in later sections to denote couplings of two processes in which the 'noise' term is common to both elements.

Definition 4.1.1 (Synchronous coupling)**.** *Consider two stochastic processes X, Y as in Example 4.1.1 with drift and diffusion coefficients b, σ and ρ, τ respectively, and driving Brownian motions W, B. We will call $(X, \tilde{Y})$ the synchronous coupling of (X, Y) if $\tilde{Y}$ satisfies*

$$d\tilde{Y}_t = \rho(\tilde{Y}_t)\,dt + \tau(\tilde{Y}_t)dW_t, \quad t > 0, \tag{4.1.2}$$

that is, if $\tilde{Y}$ is a version of Y driven by the same Brownian motion as X.

Example 4.1.2 (Ornstein-Uhlenbeck process)**.** *The Ornstein-Uhlenbeck (OU) process in $\mathbb{R}$ will repeatedly give simple, illustrative examples throughout this chapter, so we record some elementary facts here. For more on this well-loved process see for example* **?***Bakry et al. (2014).*

X is a real-valued OU process if for some $\gamma, \sigma > 0$ it satisfies the Itô stochastic differential equation for $t \geq 0$,

$$dX_t = -\gamma X_t dt + \sqrt{2\gamma\sigma^2} dB_t, \tag{4.1.3}$$

where B is a standard Brownian motion. If X_0 has the normal distribution with mean 0 and variance σ^2, then X is stationary. In that case, its covariance function is

$$K_X(s,t) \coloneqq \mathbb{E} X_s X_t = \sigma^2 e^{-\gamma|t-s|}. \tag{4.1.4}$$

An alternative characterization follows from a time change: With $A_t = e^{2\gamma t}$, the process $t \mapsto \sigma e^{-\gamma t} B_{A_t}$ for $t > -\infty$ also is a stationary OU process with covariance function (4.1.4).

We will need a more general definition of the Wasserstein distances first considered in Chapter 3, since in places we will consider distances between laws of functions rather than random variables in $\mathbb{R}^d$.

Definition 4.1.2 (Wasserstein distance on Polish spaces)**.** *We will use the following notation to mark the Wasserstein distance between probability measures on a Polish space $\mathcal{S}$ with respect to a cost function $c \in \mathcal{C}\left(\mathcal{S}^2, \mathbb{R}\right)$ Villani, C. (2009, Ch. 6),*

$$W_c\left(\mu, \nu\right) = \inf \left\{ \int_{\mathcal{S}^2} c(x,y)\, \gamma(dx, dy) \ : \quad \gamma \textit{ couples } \mu, \nu \right\}. \tag{4.1.5}$$

When there is no risk of confusion, we will use the shorthand $W_c\left(Z_1, Z_2\right)$ to denote the Wasserstein distance between the laws of random variables Z_1, Z_2.

Behind each of the propositions in this section is the following elementary result from optimal transport theory in one dimension. Consider Z_1, Z_2 real-valued random variables and cost function $c(x,y) = \ell(x-y)$ for a convex function ℓ such that $|W_c(Z_1, Z_2)| < \infty$. Write F for the CDF of Z_1 and G for the CDf of Z_2. Then the quantile coupling, also called the monotone coupling, is optimal:

$$W_c(Z_1, Z_2) = \mathbb{E}\, \ell\left(Z_1 - G^{-1}(F(Z_1))\right). \tag{4.1.6}$$

In the case where ℓ is concave and satisfies the integrability conditions for W_c, the anti-monotone or monotone decreasing coupling $(Z_1, G^{-1}(1 - F(Z_1))$ is optimal. See e.g. Santambrogio, F. (2015, Theorem 2.9 and subsequent remarks).

In particular, these results applied to the costs $|x-y|^2$ and $-|x-y|^2$ when Z_1, Z_2 are square-integrable show that $\mathbb{E}\bar{Z}_1\bar{Z}_2$ is maximized over couplings $(\bar{Z}_1, \bar{Z}_2)$ by the monotone coupling, and minimized by the anti-monotone coupling.

4.1.2 Multifidelity and surrogate modeling

We consider the setup from Pipiras and Brown (2019): Consider two continuous stochastic processes X_t, Y_t for $t \in [0,T]$ and their maxima over this time interval,

$$M^{(1)} = \max_{t\in[0,T]} X_t, \qquad M^{(2)} = \max_{t\in[0,T]} Y_t. \tag{4.1.7}$$

In this example, $M^{(1)}$ represents the maximum of a 'high-fidelity' model that is difficult to collect data on or to sample. The intense computational cost of physics based models can be a major barrier to their use in estimating rare event quantities, such as tail probabilities of the maxima, which have insufficient data for good statistical estimation. Y here is a 'low-fidelity' version of the process X for which ample data is available to estimate quantities related to its maximum $M^{(2)}$. Typically Y is constructed using procedural model-reduction techniques or via physics-based arguments. See for example Willcox (2020) for a recent conference presentation on these topics from an applied perspective.

Suppose the goal is to estimate $\mathbb{P}\left(M^{(1)} > x\right)$ for some large level x. A simple 'multifidelity' type estimator for this quantity could be

$$\hat{p}_{MF} = \hat{p}^{(2)} + \hat{p}^{(1,2)}, \tag{4.1.8}$$

where $\hat{p}^{(2)}$ is an estimate of $\mathbb{P}\left(M^{(2)} > x\right)$ based on a large number of samples from $M^{(2)}$ and $\hat{p}^{(1,2)}$ is an estimate of $\mathbb{P}\left(M^{(2)} > x\right) - \mathbb{P}\left(M^{(1)} > x\right)$ based on a smaller number of samples from $(M^{(2)}, M^{(1)})$ constrained by the difficulty of sampling from the high-fidelity model.

In Pipiras and Brown (2019), the emphasis is on classical statistical approaches to these estimates. Instead of discussing that approach, we briefly make a heuristic connection to the topics of this chapter by viewing $\hat{p}^{(1,2)}$ as a form of cost between the laws of $M^{(1)}, M^{(2)}$.

For a fixed x, define the cost $c(m_1, m_2) = \mathbb{1}\{m_1 > x, m_2 \leq x \quad \text{or} \quad m_1 \leq x, m_2 > x\}$. Then

$$\begin{aligned}\hat{p}^{(1,2)} &= \mathbb{P}\left(M^{(2)} > x\right) - \mathbb{P}\left(M^{(1)} > x\right) \\ &\leq \left|\mathbb{P}\left(M^{(2)} > x, M^{(1)} \leq x\right) - \mathbb{P}\left(M^{(2)} \leq x, M^{(1)} > x\right)\right| \\ &\leq \mathbb{E}\left(c\left(M^{(2)}, M^{(1)}\right)\right). \end{aligned} \tag{4.1.9}$$

This is stated for $M^{(2)}, M^{(1)}$ but holds for any coupling of those random variables. Since the left-hand side contains only marginal information, the optimal transport problem with cost c is an upper bound on the approximation error $\hat{p}^{(1,2)}$.

Define (M_1, M_2) to be the monotone coupling of $(M^{(1)}, M^{(2)})$ as in (4.1.6). Then if (M_1', M_2') is any other coupling,

$$\begin{aligned}\mathbb{P}(M_1 > x, M_2 \leq x) + \mathbb{P}(M_1 \leq x, M_2 \leq x) &= \mathbb{P}(M_1 \leq x) \\ &= \mathbb{P}\left(M_1' > x, M_2' \leq x\right) + P\left(M_1' \leq x, M_2' \leq x\right), \end{aligned} \tag{4.1.10}$$

and

$$\mathbb{P}(M_1 \leq x, M_2 \leq x) = F(x) \wedge G(x) \geq P(M_1' \leq x, M_2' \leq x), \tag{4.1.11}$$

by definition of the monotone coupling, where F, G are the CDFs of $M^{(1)}, M^{(2)}$ respectively Santambrogio, F. (2015, Ch. 2). Now (4.1.10), (4.1.11) show

$$\mathbb{P}(M_1 > x, M_2 \leq x) \leq \mathbb{P}(M_1 > x, M_2 \leq x). \tag{4.1.12}$$

Interchanging roles of the random variables also shows $\mathbb{P}(M_1 \leq x, M_2 > x) \leq \mathbb{P}(M_1' \leq x, M_2' > x)$. Therefore,

$$\mathbb{E}(c(M_1, M_2)) = \inf\left\{\mathbb{E}\left(c\left(M_1', M_2'\right)\right) : \quad (M_1', M_2') \text{ couples } (M^{(1)}, M^{(2)})\right\}. \tag{4.1.13}$$

In other words, the monotone coupling is optimal for cost c and thus provides the best upper bound. This is a vapid observation if $M^{(1)}, M^{(2)}$ have the same law but a useful heuristic in the more typical case when they will not.

This calculation is not too useful in practice, since we have assumed that the law of $M^{(2)}$ is not well estimated. It does, however, support the intuitive notion that we should seek to approximate the 'high-fidelity' process X with a low-fidelity version $\tilde{Y}$ that is as positively dependent as possible. In a sense, this is what the synchronous coupling in Definition 4.1.1 provides in the context of diffusions, and exploring these connections is the purpose of this chapter. This is only a heuristic, as the statements above in fact do not involve the laws of the underlying processes explicitly.

Some work in the engineering and applied mathematics literature has already started to investigate these types of questions. We make particular mention of Arbabi and Sapsis (2020) since it illustrates some of the considerations at play: There, the authors suppose we have data about a stationary stochastic dynamical system X (in their case $X_t \in \mathbb{R}^d$), which they use to create a deterministic map of convenient form transporting the data distribution to a reference stochastic differential equation Y which is stationary. Thus $T(Y)$ for some deterministic map T acts as a surrogate for X from which they can generate arbitrary amounts of data. This map, the Knothe-Rosenblatt rearrangement, is monotone and therefore preserves positive dependence, though it is not exactly a multidimensional analog of the montone coupling in one dimension Santambrogio, F. (2015, Thm. 2.23).

As we will see again in discussion following Example 4.3.1, this construction does not preserve the covariance structure of X. To account for this, the authors perform an additional optimization to match the covariance function of Y to that of X. However, it is unclear the degree to which this two-step procedure in fact gives an approximation of the studied process X that is optimal in terms of a cost of the form (4.1.5).

4.2 Couplings to approximate paths

Here we explore couplings of two stochastic processes X, Y indexed by the time interval $[0, 1]$. Therefore we consider costs in the Definition 4.1.2 that are functionals of their entire paths. In fact we study only a very tractable cost of this type,

$$c(x,y) = \int_0^1 |x_s - y_s|^2 ds. \tag{4.2.1}$$

We will look for these couplings by first expanding X, Y in bases on $L^2([0,1], ds)$, then couple their Fourier coefficients as random variables. It is important to note that we are not interested in the case where Y is a lower-dimensional version of X. In that case, one could just truncate the expansion to a desired dimensionality to achieve optimality for the transport problem in Definition 4.1.2 with cost (4.2.1). That result is a standard one from analysis but is also implied by Proposition 4.2.2.

We draw inspiration in this approach from Banerjee and Kendall (2016), which constructs a maximal coupling between a Brownian motion and its path integral by similar means, where maximality is in the sense that it minimizes at each time the probability two versions of the process have not yet collided. The basis-coefficient coupling we consider here is much simpler only because the problem of optimizing (4.2.1) is much simpler than that of finding a maximal coupling.

Finally, it bears mention that there are a number of substantial optimal transport results from Wiener space analysis, but these are rather too abstract for our perspective here. We refer the reader to Feyel and Üstünel (2004); Fang et al. (2010).

4.2.1 Basis coefficient couplings

The following is a statement of Mercer's Theorem Lax (2002, Ch. 30.5, Theorem 11), with a formulation for covariance kernels of stochastic processes.

Theorem 4.2.1. *Consider a continuous function $K(s,t)$ and define a kernel operator K : $L^2([0,1], ds) \mapsto L^2([0,1], ds)$ as*

$$Kg(s) = \int_0^1 K(s,t) g(t) dt, \tag{4.2.2}$$

and suppose $(Kg, g) \geq 0$ for all $g \in L^2([0,1], ds)$, where (f,g) is the standard inner product. Then, there exists a summable sequence of non-negative real numbers $\lambda_1 \geq \lambda_2 \ldots$ and an *orthonormal basis of eigenvectors* $\{e_i\}_{i\geq 1}$ *such that*

$$K(s,t) = \sum_{1}^{\infty} \lambda_i e_i(s) e_i(t). \tag{4.2.3}$$

In particular, for a continuous stochastic process X_t, $t \in [0,1]$ *with mean process* $\mathbb{E}X_s = \mu_s^X$ *and* $s \mapsto \mathbb{E}X_s^2$ *is continuous for all* s*, then by Fubini's theorem the covariance function* $K_X(s,t) := \mathbb{E}(X_s - \mu_s^X)(X_t - \mu_t^X)$ *satisfies* $(K_X g, g) \geq 0$ *and has an expansion of the form* (4.2.3). *We may expand* $X - \mu^X$ *as an element of* $L^2([0,1], ds)$ *in the basis* $\{e_i^X\}_{i\geq 1}$ *to write*

$$X_t = \mu_t^X + \sum_{1}^{\infty} \sqrt{\lambda_i^X}\, Z_i^X\, e_i^X(t), \tag{4.2.4}$$

where $\{Z_i^X\}_{i\geq 1}$ *is such that* $\mathbb{E}Z_i^X Z_j^X = \mathbb{1}_{i=j}$ *and* $\mathbb{E}Z_i^X = 0$ *for all* i, j. *In* (4.2.4) *if* $\int_0^1 X_s e_i^X(s) ds = 0$ *a.s. then* $\sqrt{\lambda_i^X} Z_i^X = 0$ *by definition.*

The following proposition observes that the optimal transport problem in squared integral cost is achieved by applying the monotone coupling or anti-monotone coupling according to whether their corresponding basis elements are positively or negatively correlated in L^2 — at least in the case where the coefficients form an independent sequence of random variables.

Proposition 4.2.2. *Consider stochastic processes* X, Y *such that conditions leading to* (4.2.4) *hold, and further assume their coefficients* Z^X, Z^Y *each are sequences of independent random variables. With the cost function* $c : L^2 \times L^2 \mapsto \mathbb{R}_+$ *given as* $c(f,g) = \int_0^1 |f(s) - g(s)|^2\, ds$, *recalling Definition 4.1.2, it holds that*

$$W_c(X,Y) = \mathbb{E}\int_0^1 |X_s - \tilde{Y}_s|^2\, ds, \tag{4.2.5}$$

where $\tilde{Y}$ *is defined as follows. Write* F_i *for the CDF of* Z_i^X *and* G_i *for that of* Z_i^Y. *Set*

$$\tilde{Z}_i^Y = \begin{cases} G_i^{-1}(F_i(Z_i^X)) & \text{if } (e_i^X, e_i^Y) \geq 0, \\ G_i^{-1}(1 - F(Z_i^X)) & \text{otherwise,} \end{cases} \tag{4.2.6}$$

and

$$\tilde{Y}_t := \mu_t^Y + \sum_1^{\infty} \sqrt{\lambda_i^Y}\, \tilde{Z}_i^Y\, e_i^Y(t), \quad t \in [0,1]. \tag{4.2.7}$$

This implies $\tilde{Y} \stackrel{d}{=} Y$ *as an element of* $L^2([0,1], ds)$.

Example 4.2.1 (Gaussian processes)**.** *In the case where* X, Y *are Gaussian processes, the line following* (4.2.4) *implies* Z^X, Z^Y *are i.i.d. standard normal random variables. Thus* (4.2.6) *reduces to*

$$\tilde{Z}_i^Y = \begin{cases} Z_i^X & \text{if } (e_i^X, e_i^Y) \geq 0, \\ -Z_i^X & \text{otherwise} \end{cases} \tag{4.2.8}$$

Remark 4.2.1 (Computational perspective)**.** *Suppose now that we have data from independent samples of the processes* X, Y *at discrete time points* $0 \leq t_1 \leq \ldots t_N$.

We now can estimate the basis expansion in (4.2.4) *by performing principle component analysis on the matrix of sample covariances at each pair of time points Ramsay and Silverman (2005, e.g.), from which estimates of the coefficients* Z^Y, Z^X *can be computed. In principle, the data could also be used to explicitly construct a sample-based* $\tilde{Z}^Y$ *from* (4.2.6) *and therefore* $\tilde{Y}$*, at least in cases where the independence assumptions for* Z^X, Z^Y *are reasonable. For a discussion of the general case see Remark 4.2.2.*

Finally, we note that there is a large body of work in computational mathematics fields on using spectral decomposition methods to construct surrogates for dynamical systems. Since that is not a focus of this work, we simply note the connection, list some references and leave development of Proposition 4.2.2 for computational applications to future work. To name just a few, see Kersaudy et al. (2015); Xiu and Karniadakis (2002); Soize and Ghanem (2004) and references therein.

Remark 4.2.2 (Dependent coefficients)**.** *We have stated Proposition 4.2.2 in the case where* Z^X *and* Z^Y *each are independent sequences because, in the general case, we must somehow define the joint distribution of* $\tilde{Z}^Y$ *from* (4.2.7) *so that* $\tilde{Y}$ *is equal in distribution to* Y.

The primary value of Proposition 4.2.2 is that it shows a somewhat thorny optimal transport problem (4.2.5) *can be solved explicitly via solutions to a sequence of much simpler problems,* (4.2.6). *In the general case, we could attempt to reformulate* (4.2.5) *as an optimal transport problem between* Z^Y, Z^X *but it is not yet clear why that would be any simpler.*

However, consider now the setup of Remark 4.2.1. This procedure could be used to generate histograms for the sample distributions of Z^X, Z^Y*. Then, we could formulate an appropriate an optimal transport problem for these sample distributions, which could be solved via standard methods Flamary and Courty (2017); Peyré and Cuturi (2018, e.g.). This opens the door to solving* (4.2.5) *from a computational perspective in more general scenarios, and to solving similar problems with different path costs* c *that are amenable to reformulation as transport problems for the basis coefficient vectors. Determining whether such a procedure is indeed a reasonable approach to applied problems in multifidelity and surrogate modeling is part of this author's plans for future work.*

We close this section with a result from Bion-Nadal and Talay (2019, Section 2.1, p 1620-1621) that investigates the problem (4.2.5) by looking for solutions over a more limited, but reasonable, class of couplings. It shows that, when considering only couplings adapted to the same filtration, the synchronous coupling minimizes the L^2 Wasserstein distance given in (4.2.5). As such it is a kind of bridge between this section and the next.

Example 4.2.2 (Synchronous coupling optimality). *Consider* $X, Y \in \mathcal{C}([0,\infty), \mathbb{R})$ *as two solutions of the type given in Example* (4.1.1)*, with drift and diffusion coefficients* b, σ *and* ρ, τ *respectively. Let* W *be the driving Brownian motion of* X *and* B *that of* Y*, which are constructed on the same probability space.*

Define $\mathcal{A}$ *to be the class of couplings* $(X, \bar{Y})$ *such that* $\tilde{Y}$ *satisfies*

$$d\bar{Y}_t = \rho(\bar{Y}_t)dt + \tau(\bar{Y}_t)\left(u_t dW_t + \sqrt{1-u_t^2}d\bar{B}_t\right), \tag{4.2.9}$$

where W *is the driving Brownian motion of* X*,* $\bar{B}$ *is an independent Brownian motion on the same probability space, and* u *is a stochastic control process adapted to the joint filtration of* (W, B) *such that* $u_t \in [-1,1]$ *for* $t \geq 0$*. Thus* u *gives the degree to which the coupling is synchronous.*

Then, the synchronous coupling $(X, \tilde{Y})$ with $u_t = 1$ for all $t \geq 0$ satisfies

$$\inf_{(X,\bar{Y})\in\mathcal{A}} \mathbb{E}\int_0^1 |X_s - \bar{Y}_s|^2\, ds = \mathbb{E}\int_0^1 |X_s - \tilde{Y}_s|^2\, ds, \tag{4.2.10}$$

A few comments on this example: First is that the main result, Theorem 2.2, of Bion-Nadal and Talay (2019) is to extend the solution given in (4.2.10) to multidimensional diffusionsi via stochastic optimal control methods. There, however, the analogous process to u is not explicit. Second, we note Kendall (2009); Émery (2005) imply the class $\mathcal{A}$ of solutions considered is identical to the class of co-adapted couplings, that is couplings $(\bar{X}, \bar{Y})$ such that the Markov property holds marginally for $\bar{X}$ conditional on the *joint history* of $(\bar{X}, \bar{Y})$ not just on the history of $\bar{X}$ — and the same for $\bar{Y}$.

At present, the relationship between solutions to problem (4.2.10) and (4.2.5) is unclear, except that since the former optimizes over a smaller class of models the optimal value achieved must be greater than or equal to the optimal value of the second. More interesting is to ask whether the synchronous coupling also provides a solution (though perhaps not the unique one) to (4.2.5) as well.

Though we do not yet have a satisfying answer to this question, we close with a discussion showing that in the Gaussian case (4.2.7) represents a deterministic operation on the basis elements of the Wiener chaos expansion of X. While this does not mark the optimal coupling $(X, \tilde{Y})$ as synchronous, in the sense of Example 4.2.2 or Definition 4.1.1, it does show $\tilde{Y}$ is a functional of W, the driving Brownian motion of X. That fact is evident from the construction in 4.2.7 since the Z_i^X themselves are of functionals of X and hence of W, but here we make this more explicit.

We recall only the bare minimum of facts from Wiener space analysis needed, as given in Nualart (2006, Ch. 1) and specialized to the scenario of Example 4.2.1.

Consider $W_s, s \in [0, 1]$ a standard Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathcal{F}$ is taken to be the filtration generated by W. Write $L^2([0, 1] \times \Omega)$ for the $\mathcal{F}$-measurable processes V such that $\mathbb{E}\int_0^1 V_s^2 ds < \infty$. Then V has the decomposition

$$V_s = \sum_0^\infty I_n(f_n(\cdot, s)), \tag{4.2.11}$$

where $I_n(f_n(\cdot,s)) = n! \int_0^1 \int_0^{u_n} \dots \int_0^{u_1} f_n(u_1, u_2 \dots u_n, s)\, dW_{u_1} \dots dW_{u_n}$ is an iterated Itô integral. The functions $\{f_n(u,s)\}_{n\geq 1}$ are symmetric in u, are uniquely determined by V, and $f_n \in L^2([0,1]^{n+1}, ds)$ for each n.

In addition, the following orthogonality and isometry properties hold: For each $m, n \geq 0$ and symmetric functions $g \in L^2([0,1]^m, ds), h \in L^2([0,1]^n, ds)$

$$\mathbb{E} I_m(g) I_n(h) = \begin{cases} 0 & \text{if } m \neq n \\ n!(g,h) & \text{if } m = n, \end{cases} \tag{4.2.12}$$

where $(\cdot,\cdot)$ is the standard inner product on $L^2([0,1]^n, ds)$. As a result, writing $\|\cdot\|_{n+1}^2$ for the standard norm in $L^2([0,1]^{n+1})$,

$$\mathbb{E} \int_0^1 V_s^2 ds = \sum_0^\infty n! \|f_n\|_{n+1}^2. \tag{4.2.13}$$

Now consider $X, \tilde{Y}$ from Example 4.2.1 and without loss of generality assume $EX_s = EY_s = 0$ for all $s \in [0,1]$. From (4.2.4) it follows that $Z_i^X = \frac{1}{\sqrt{\lambda_i^X}} \int_0^1 X_s e_i^X(s) ds$ for each i. Therefore, by definition $\tilde{Y}$ is measurable with respect to the filtration generated by W and $\tilde{Y} \in L^2([0,1] \times \Omega)$. Therefore, both $\tilde{Y}$ and X have expansions of the form (4.2.11),

$$X_s = \sum_0^\infty I_n(f_n(\cdot,s)), \qquad \tilde{Y}_s = \sum_0^\infty I_n(g_n(\cdot,s)), \tag{4.2.14}$$

First note that since $Z_i^Y = \frac{1}{\sqrt{\lambda_i^Y}} \int_0^1 e_i^Y(s) Y_s ds$ are standard normal we have $\lambda_i^Y > 0$ for all i, with the same holding for λ_i^X. Define the operator $A : \mathcal{D}(A) \mapsto L^2([0,1], ds)$ — with $\mathcal{D}(A) = \{g \in L^2 : \|A(g)\|^2 < \infty\}$ — as

$$(Ag)(s) = \sum_1^\infty s_i \sqrt{\frac{\lambda_i^Y}{\lambda_i^X}} \left(e_i^X, g\right) e_i^Y(s), \quad s_i = \operatorname{sign}(e_i^X, e_i^Y), \tag{4.2.15}$$

where e^X, e^Y are the basis functions for X, Y from (4.2.4).

If it holds that $f_n(u,\cdot) \in \mathcal{D}(A)$, we can show

$$\tilde{g}_n(u,s) \coloneqq A(f_n(u,\cdot))(s) = g_n(u,s), \quad \text{for all } u \in [0,1]^n, s \in [0,1]. \tag{4.2.16}$$

Since this transformation preserves symmetry in the u variable, by the uniqueness of the expansion (4.2.14) we need only show that for any square-integrable random variable F measurable with respect to the σ-field generated by W, for all $i \geq 1, n \geq 0$ it holds

$$\begin{aligned}
\mathbb{E}\left[F\int_0^1 e_i^Y(s)I_n(\tilde{g}_n(\cdot,s))ds\right] &= s_i\sqrt{\frac{\lambda_i^Y}{\lambda_i^X}}\int_0^1 e_i^X(s)\mathbb{E}\left[F\int_0^1 e_i^X(s)I_n(f_n(\cdot,s))ds\right] \\
&= s_i\sqrt{\lambda_i^Y}\mathbb{E}\left[FZ_i^X\right] = \sqrt{\lambda_i^Y}\mathbb{E}\left[F\tilde{Z}_i^Y\right]. \qquad (4.2.17)
\end{aligned}$$

The proof of (4.2.17) follows from observing that with $F = \sum_{m\geq 0} I_m(h_m)$, we have by (4.2.12)

$$\begin{aligned}
\mathbb{E}F\int_0^1 e_i^Y(s)I_n(\tilde{g}_n(\cdot,s))ds &= \int_0^1 e_i^Y(s)\sum_{m\geq 0} m!(h_m,\tilde{g}_m(\cdot,s))_{L^2([0,1]^m)}\,ds && (4.2.18)\\
&= \sum_{m\geq 0} m!\int_{[0,1]^m} h_m(u)(e_i^Y,\tilde{g}_m(u,\cdot))\,du && (4.2.19)\\
&= \sum_{m\geq 0} m!\int_{[0,1]^m} h_m(u)s_i\sqrt{\frac{\lambda_i^Y}{\lambda_i^X}}(e_i^X,f_m(u,\cdot))\,du && (4.2.20)\\
&= s_i\sqrt{\frac{\lambda_i^Y}{\lambda_i^X}}\mathbb{E}F\int_0^1 e_i^X(s)I_n(f_n(\cdot,s))ds. && (4.2.21)
\end{aligned}$$

(4.2.16) follows. Again, we note this requires $f_n(u,\cdot) \in \mathcal{D}(A)$, which at present is assumed and does not immediately follow to the best of our knowledge.

4.3 Monotone coupling as a synchronously coupled process

The following proposition is a simple consequence of the Fokker-Planck equation, Proposition 2.4 of Alfonsi et al. (2014), Itô's formula and generic optimal transport results in one dimension. We defer the proof to Section 4.4.

In what follows, $\partial_u g(f(x))$ denotes $\partial_u g(u)$ evaluated at $u = f(x)$ for functions f, g, and to make a clear distinction we write $\partial_x g \circ f(x)$ instead of $\partial_x g(f(x))$. Similarly, we write $\partial_t g_t(f_t(x))$ for $\partial_t g_t(u)$ at $u = f_t(x)$ as distinct from $\partial_t\left(g_t \circ f_t(x)\right) = \partial_t g_t(f_t(x)) + \partial_u g_t(f_t(x))\,\partial_t f_t(x)$.

Proposition 4.3.1. *Suppose the processes X, Y satisfy Assumption 4.1.1 with respective diffusion, drift coefficients σ, b and τ, ρ and driving Brownian motions W, B. Without loss of generality we may take these to be constructed on a common probability space.*

Fix $T_1 > T_0 > 0$. Then, for any cost function of the form $c(x,y) = \ell(x-y)$ with ℓ convex such that $\sup_{t\in[T_0,T_1]} |W_c(X_t, Y_t)| < \infty$, it holds

$$\mathbb{E}\,\ell\left(X_t - \tilde{Y}_t\right) = W_c(X_t, Y_t), \quad \text{for } t \in [T_0, T_1], \tag{4.3.1}$$

where $\tilde{Y}_t = G_t^{-1}(F_t(X_t))$ for $t \in [T_0, T_1]$ satisfies

$$d\tilde{Y}_t = \tilde{\rho}(t, X_t)dt + \tau\left(G_t^{-1}(F_t(X_t))\right)\tilde{r}(t, X_t)\, dW_t, \tag{4.3.2}$$

with W the driving Brownian motion of X and time-varying coefficients,

$$\begin{aligned}
\tilde{r}(t,x) &= \frac{\sigma(x)}{\tau\left(G_t^{-1}(F_t(x))\right)}\partial_x G_t^{-1} \circ F_t(x) && (4.3.3)\\
\tilde{\rho}(t,x) &= 2\partial_t\left(G_t^{-1} \circ F_t(x)\right) + \left(\tilde{r}^2(t,x) - 1\right)\partial_t G_t^{-1}(F_t(x))\\
&\quad + \partial_x G_t^{-1} \circ F_t(x)\left(2b(x) - \frac{1}{2}\partial_x \sigma^2(x)\right)\\
&\quad - \tilde{r}^2(t,x)\left(\rho\left(G_t^{-1}(F_t(x))\right) - \frac{1}{2}\partial_y \tau^2\left(G_t^{-1}(F_t(x))\right)\right) && (4.3.4)
\end{aligned}$$

Again we emphasize the perspective of using one system to approximate the other, not to answer convergence questions as in the Chapters 2, 3, in which the proposition would make no sense.

Remark 4.3.1. *In general the proposition cannot be extended to define $\tilde{Y}_t$ as a process starting from $t = 0$ because of possible degeneracy in the initial condition for the transition density. Consider for example X a standard Brownian motion started from a point x_0 and Y any process satisfying the assumptions that is stationary with respect to the standard Normal distribution.*

Then for $t > 0$, $\tilde{Y}_t = \frac{X_t - x_0}{\sqrt{t}}$, which diverges as $t \to 0$ and thus the process is not well-defined when started at time $t = 0$. The CDF of X is of course not differentiable at $t = 0$, and Itô's theorem cannot be applied to $F_t(X_t)$.

Though this problem arises in general when X, Y have degenerate initial distributions, there are a few reasonable ways to circumvent it. The simplest is to consider the processes as starting from a small fixed time, as in the next example.

Example 4.3.1 (Quantile coupling does not couple paths)**.** *Though $(X_t, \tilde{Y}_t)$ of Proposition 4.3.1 couples (X_t, Y_t) for each time t, this in general does not provide a coupling of the processes (X, Y) as random variables in $\mathcal{C}([T_0, T_1], \mathbb{R})$. The first clue is that the coefficients of $\tilde{Y}$ are time-variant whereas those of Y are not. However, the OU process once again provides a helpful illustration.*

Consider for example the case where both X and Y are stationary OU processes in time-changed form, with $X_t = e^{-\gamma t} W_{A_t^\gamma}$ and $Y_t = e^{-t} B_{A_t}$, where $\gamma > 0$, $A_t^\gamma = e^{2\gamma t}$ and $A_t = A_t^1$ in this notation.

Then both X_t, Y_t have the standard normal distribution, so that $\tilde{Y}_t = X_t$ for all t. However, Y and $\tilde{Y}$ do not have the same distribution as stochastic processes since $\mathbb{E} Y_t Y_s \neq \mathbb{E} \tilde{Y}_t \tilde{Y}_s$.

This example also shows clearly that Proposition 4.3.1 is only meaningful in cases where either X or Y is non-stationary. If both were stationary, the monotone coupling at each time point produces a cost (4.3.1) that is constant in time. In addition, since $\tilde{Y}$ is not equal in distribution as a process to Y by the previous example, the proposition does not provide a useful means for approximating X with $\tilde{Y}$ if one is interested in path-dependent quantities such as maxima or time integrals. This is in contrast to the couplings discussed in Section 4.2, which do couple the processes and not just the time marginals.

4.4 Proofs

Proof of 4.2.2. Consider any coupling $(\bar{X}, \bar{Y})$ of (X, Y) each term of which ncessarily has an expansion of the form (4.2.4) with respect to the bases $\{e_i^Y\}_{i \geq 1}, \{e_i^X\}_{i \geq 1}$. Expanding the terms in the cost function, we see

$$\mathbb{E} \int_0^1 |\bar{X}_s - \bar{Y}_s|^2 ds = \mathbb{E} \int_0^1 | \left(\bar{X}_s - \mu_s^X \right) - \left(\bar{Y}_s - \mu_s^Y \right) |^2 \, ds - \int_0^1 |\mu_s^X - \mu_s^Y|^2 ds. \tag{4.4.1}$$

Since the first term is the only one to be optimized, without loss of generality we assume $\mu_s^X = \mu_s^Y = 0$ for all s. Since Z^X, Z^Y each are independent sequences of centered and scaled random variables, to couple X, Y we need only couple (Z_i^X, Z_i^Y) term by term. Write these couplings as $(\bar{Z}_i^X, \bar{Z}_i^Y)$. Then,

$$\mathbb{E}\int_0^1 |\bar{X}_s - \bar{Y}_s|^2 ds = \mathbb{E}\int_0^1 |X_s|^2\, ds + \mathbb{E}\int_0^1 |Y_s|^2\, ds \\ - 2\sum_1^{\infty} \sqrt{\lambda_i^Y \lambda_i^X}(e_i^X, e_i^Y)\mathbb{E}\bar{Z}_i^X \bar{Z}_i^Y, \tag{4.4.2}$$

where we have used the fact that $\mathbb{E}\bar{Z}_i^X \bar{Z}_j^Y = 0$ for $i \neq j$. We maximize each term in the sum in the second line of (4.4.2) using elementary optimal transport theory in one dimension. Write F_i for the CDF of Z_i^X (so also that of $\bar{Z}_i^X$) and G_i for that of Z_i^Y. By (4.1.6) and the subsequent discussion, we know

$$\mathbb{E}Z_i^X G_i^{-1}(1 - F_i(Z_i^X)) \leq \mathbb{E}\bar{Z}_i^X \bar{Z}_i^Y \leq \mathbb{E}Z_i^X G_i^{-1}(F_i(Z_i^X)). \tag{4.4.3}$$

The proposition now follows by applying (4.4.3) and the definition (4.2.6) of $\tilde{Z}_i^Y$ to (4.4.2). $\square$

Proposition 4.3.1 will follow from Alfonsi et al. (2014, Proposition 2.4) and a statement about regularity of $(t, x) \mapsto F_t(x)$ in its proof, which is a consequence of the regularity of transition densities for solutions such as those in 4.1.1. For reference we state the latter result here first as a lemma.

A minor technical comment: Alfonsi et al. (2014) imposes additional regularity on the second-order derivatives of σ^2. However, the cited reference on existence and regularity of transition densities needed for this Lemma shows the assumptions in 4.1.1, namely bounded continuous second-order derivatives, are sufficient.

Lemma 4.4.1. *Suppose X is a diffusion as in 4.1.1, in particular such that Assumptions 4.1.1 hold. Fix a time $T > 0$. Writing $F_t(x)$ for the CDf of X_t at $t > 0$, then for each $t \in (0, T]$ the function $x \mapsto F_t(x)$ is invertible, and $\partial_x F_t(x), \partial_u F_t^{-1}(u) > 0$ for $t \in (0, T], u \in (0, 1), x \in \mathbb{R}$. The function $(t, x) \mapsto F_t(x)$ is $\mathcal{C}^{1,2}$ on $(0, T] \times \mathbb{R}$ and the inverse CDF $(t, u) \mapsto F_t^{-1}(u)$ is $\mathcal{C}^{1,2}$ on $(0, T] \times (0, 1)$. Furthermore, they satisfy the following partial differential equations,*

$$\partial_t F_t(x) = \frac{1}{2}\partial_x\left(\sigma^2(x)\partial_x F_t(x)\right) - b(x)\partial_x F_t(x), \tag{4.4.4}$$

$$\partial_t F_t^{-1}(u) = -\frac{1}{2}\partial_u \frac{\sigma^2\left(F_t^{-1}(u)\right)}{\partial_u F_t^{-1}(u)} + b\left(F_t^{-1}(u)\right) \tag{4.4.5}$$

Proposition 4.3.1 now follows from Lemma 4.4.1 and direct application of Itô's theorem to the processes started from T_0.

Proof of Proposition 4.3.1. By (4.1.6), to prove (4.3.1) we need only show the process $G_t^{-1}(F_t(X_t))$ exists for $t \in [T_0, T_1]$ and has the representation (4.3.2). Write $U_t = F_t(X_t)$ for $t \in [T_0, T_1]$. By Lemma 4.4.1 and Itô's formula we have

$$\begin{aligned} dU_t &= \partial_t F_t(X_t)dt + \partial_x F_t(X_t)dX_t + \frac{\sigma^2(X_t)}{2}\partial_{xx}F_t(X_t)dt \\ &= \left[\partial_x F_t(X_t)\left(2b(X_t) - \frac{1}{2}\partial_x \sigma^2(X_t)\right) + 2\partial_t F_t(X_t)\right]dt + \partial_x F_t(X_t)\sigma(X_t)dW_t \\ &= \tilde{b}(t, X_t)dt + \tilde{\sigma}(t, X_t)dW_t. \end{aligned} \tag{4.4.6}$$

Note $U_t \in (0,1)$ for $t \in [T_0, T_1]$ a.s.: Lemma 4.4.1 states $\partial_x F_t(x)$ is strictly positive for all such t and $x \in \mathbb{R}$, so $U_t = 0, 1$ could occur only if X diverges in finite time, which is not the case for diffusions of the type in Example 4.1.1 Le Gall (2013, Thm. 5.16 for standard moment bounds, e.g.). Applying (4.4.5) to $G_t^{-1}(x)$ with the drift, diffusion coefficients ρ, τ of Y, then rearranging terms,

$$\frac{1}{2}\partial_{uu}G_t^{-1}(u) = \frac{\left(\partial_{uu}G_t^{-1}(u)\right)^2}{\tau^2(G_t^{-1}(u))}\left[\partial_t G_t^{-1}(u) + \frac{1}{2}\partial_y \tau^2(G_t^{-1}(u)) - \rho(G_t^{-1}(u))\right], \tag{4.4.7}$$

where we recall the ellipticity Assumption 4.1.1 that τ^2 is uniformly bounded away from zero. Substituting (4.4.7) in the statment of Itô's theorem again,

$$
\begin{aligned}
dG_t^{-1}(U_t) = \partial_t G_t^{-1}(U_t)dt + \partial_u G_t^{-1}(U_t)dU_t& \\
+ \tilde{\sigma}^2(t,X_t)\left(\frac{\partial_{uu}G_t^{-1}(U_t)}{\tau(G_t^{-1}(U_t)}\right)^2 &\left[\partial_t G_t^{-1}(u) + \frac{1}{2}\partial_y \tau^2(G_t^{-1}(u)) - \rho(G_t^{-1}(u))\right] \\
= \partial_u G_t^{-1}(U_t)\tilde{\sigma}(t,X_t)\, dW_t &+ 2\left(\partial_t G_t^{-1}(U_t) + \partial_u G_t^{-1}(U_t)\partial_t F_t(X_t)\right) \\
- \partial_t G_t^{-1}(U_t) &+ \tilde{\sigma}^2(t,X_t)\left(\frac{\partial_{uu}G_t^{-1}(U_t)}{\tau(G_t^{-1}(U_t))}\right)^2 \partial_t G_t^{-1}(u) \\
+ \partial_u G_t^{-1}(U_t)&\partial_x F_t(X_t)\left(2b(X_t) - \frac{1}{2}\partial_x \sigma^2(X_t)\right) \\
- \tilde{\sigma}^2(t,X_t)&\left(\frac{\partial_{uu}G_t^{-1}(U_t)}{\tau(G_t^{-1}(U_t))}\right)^2 \left[\rho(G_t^{-1}(u)) - \frac{1}{2}\partial_y \tau^2(G_t^{-1}(u))\right]. \qquad (4.4.8)
\end{aligned}
$$

(4.3.2) now follows from the identifications in (4.3.3), taking into consideration the notational guidance given before the proposition statement. □

APPENDIX A: HITTING TIME ESTIMATES

The following lemmas primarily support statements leading to the proof of Theorem 2.1.1. We begin with the observation that $\tau_0^H := \sigma(0)$ is finite a.s. for all initial conditions.

Lemma .0.2. *For each $(h, \nu) \in S$,*

$$\mathbb{P}_{(h,\nu)}\left(\sigma(0) < \infty\right) = 1$$

Proof. $\sigma(0) = 0$ for $(h, \nu) \in \partial S = \{0\} \times (-g/\gamma, \infty)$. For $(h, \nu) \in S^\circ$ and $t < \sigma(0)$, recall by (2.1.4),

$$H_t \leq h + \frac{\nu}{\gamma} + \frac{g}{\gamma^2} - B_t - \frac{g}{\gamma}t$$

The lemma follows by noting that the bounding process in the above inequality is a Brownian motion with negative drift starting from a positive point and hence hits zero in finite time almost surely. $\square$

Lemma .0.3. *For any $-g/\gamma < a < b$, $\ell > \frac{b-a}{\gamma a+g} > 0$ and $m \geq 1$,*

$$\sup_{\nu \in [a,b]} \mathbb{P}_{(0,\nu)}\left(\tau^V_{[a,b]^c} > m(\ell+1)\right) \leq \left[\mathbb{P}\left(B_1 \leq b - a + (1+\gamma)b + g\right)\right]^m.$$

Proof. For $t \leq \tau^V_{[a,b]^c}$ and $\nu \in [a, b]$, recalling that $L_t = \sup_{u \leq t}\left(B_u - S_u\right)$ and from (2.1.1), we have $L_t \geq B_t - S_t \geq B_t - tb$ and $b \geq V_t \geq \nu - (\gamma b + g)t + L_t$. This gives

$$B_t \leq b - \nu + ((1+\gamma)b + g)t \leq b - a + ((1+\gamma)b + g)t$$

for all $t \leq \tau^V_{[a,b]^c}$. Fix $\ell > \frac{b-a}{\gamma a+g} > 0$. Suppose $\tau^V_{[a,b]^c} \geq 1 + \ell$. Then if $\sigma(1) > 1 + \ell$, L_t must be constant on $[1, 1+\ell]$, and using (2.1.1) and (2.1.5) shows

$$V_{1+\ell} - V_1 \leq -\ell(\gamma a + g) < a - b \qquad \Longrightarrow \qquad V_{1+\ell} < a,$$

a contradiction. Thus, $\tau^V_{[a,b]^c} \geq 1+\ell$ implies $\sigma(1) \leq 1+\ell$. For any $m \geq 1$, applying the strong Markov property at $\sigma(1)$,

$$\begin{aligned}
&\sup_{\nu\in[a,b]} \mathbb{P}_{(0,\nu)}\left(\tau^V_{[a,b]^c} > m(\ell+1)\right) \\
&\leq \sup_{\nu\in[a,b]} \mathbb{E}_{(0,\nu)}\left(\mathbb{1}_{\{B_1\leq b-a+(1+\gamma)b+g,\ \sigma(1)\leq 1+\ell,\ V_{\sigma(1)}\in[a,b]\}}\mathbb{P}_{(0,V_{\sigma(1)})}\left(\tau^V_{[a,b]^c} > m(\ell+1)-(\ell+1)\right)\right) \\
&\qquad \leq \mathbb{P}\left(B_1 \leq b-a+(1+\gamma)b+g\right) \sup_{\nu\in[a,b]} \mathbb{P}_{(0,\nu)}\left(\tau^V_{[a,b]^c} > (m-1)(\ell+1)\right).
\end{aligned}$$

An induction argument gives the result. □

Lemma .0.4. *Fix $b = -\frac{g-g/2(1+\gamma)}{1+\gamma}$ and $a = -\frac{g+g/2\gamma}{1+\gamma}$. There exist positive constants c and $t'(\gamma,g)$ such that*

$$\sup_{(h,\nu)\in[0,tg/4\gamma(1+\gamma)]\times(a,b)} \mathbb{P}_{(h,\nu)}\left(\sigma(\tau^V_a)-\tau^V_a > t,\ \tau^V_a < \tau^V_b\right) \leq e^{-ct},$$

for all $t > t'(\gamma,g)$.

Proof. Lemma .0.3 shows there exists a $\tilde{p}\in(0,1)$ and $C>0$ depending on γ, g such that for $t \geq C$,

$$\sup_{\nu\in(a,b)} \mathbb{P}_{(0,\nu)}\left(t < \tau^V_a \wedge \tau^V_b\right) \leq \tilde{p}^{t/C}. \tag{.0.9}$$

Suppose $h>0$. (2.1.5) shows V cannot reach the level $b>\nu$ without the gap process H hitting zero, which implies $\sigma(0) < \tau^V_b$. In addition, (2.1.6) implies $\tau^V_a \geq \frac{1}{\gamma}\log\left(\frac{\nu+g/\gamma}{a+g/\gamma}\right)$ with equality when $\sigma(0)\geq\tau^V_a$. Therefore if $\sigma(0)\geq\tau^V_a$, $\tau^V_a\wedge\tau^V_b \leq \frac{1}{\gamma}\log\left(\frac{\nu+g/\gamma}{a+g/\gamma}\right) \leq \frac{1}{\gamma}\log\left(\frac{b+g/\gamma}{a+g/\gamma}\right)$. Hence, for $t > \frac{1}{\gamma}\log\left(\frac{b+g/\gamma}{a+g/\gamma}\right)$, we obtain

$$\begin{aligned}
\mathbb{P}_{(h,\nu)}\left(t<\tau^V_a\wedge\tau^V_b\right) &= \mathbb{P}_{(h,\nu)}\left(\sigma(0)<\tau^V_a,\ t<\tau^V_a\wedge\tau^V_b\right) \\
&= \mathbb{P}_{(h,\nu)}\left(\frac{1}{\gamma}\log\left(\frac{\nu+g/\gamma}{a+g/\gamma}\right) < \sigma(0) < \tau^V_a,\ t<\tau^V_a\wedge\tau^V_b\right) \\
&\quad + \mathbb{P}_{(h,\nu)}\left(\sigma(0)\leq\tau^V_a\wedge\frac{1}{\gamma}\log\left(\frac{\nu+g/\gamma}{a+g/\gamma}\right),\ t<\tau^V_a\wedge\tau^V_b\right) \\
&= \mathbb{P}_{(h,\nu)}\left(\sigma(0)\leq\frac{1}{\gamma}\log\left(\frac{\nu+g/\gamma}{a+g/\gamma}\right),\ t<\tau^V_a\wedge\tau^V_b\right).
\end{aligned} \tag{.0.10}$$

(2.1.5) shows that when $\sigma(0) < \tau_a^V$, $V_{\sigma(0)} \in (a,b)$. As a result, (.0.9), (.0.10) and the strong Markov property at $\sigma(0)$ show there exists positive constants c, C and $t_0(\gamma, g)$ such that that for any $(h,\nu) \in (0,\infty) \times (a,b)$ and $t > t_0(\gamma, g)$,

$$\begin{aligned}
\mathbb{P}_{(h,\nu)}\left(t < \tau_a^V \wedge \tau_b^V\right) &= \mathbb{E}_{(h,\nu)}\left(\mathbb{P}_{(0,V_{\sigma(0)})}\left(t - \sigma(0) < \tau_a^V \wedge \tau_b^V\right) \mathbb{1}_{\sigma(0) \leq \frac{1}{\gamma}\log\left(\frac{\nu+g/\gamma}{a+g/\gamma}\right)}\right) \\
&\leq \sup_{\nu \in (a,b)} \mathbb{P}_{(0,\nu)}\left(t - \frac{1}{\gamma}\log\left(\frac{b+g/\gamma}{a+g/\gamma}\right) < \tau_a^V \wedge \tau_b^V\right) \leq e^C\, e^{-ct} \leq e^{-ct/2}. \quad (.0.11)
\end{aligned}$$

Starting from $(H_0, V_0) = (0, \nu)$ for some $\nu \in (a,b)$, $\sigma(\tau_a^V) - \tau_a^V$ is large if one of the following two events happen: (i) either the gap is large at τ_a^V or (ii) the gap at τ_a^V is not large but the gap remains positive for a large time after τ_a^V. We handle these cases separately and show that in either case, the Brownian particle has to attain a large negative value before τ_b^V, leading to exponentially small probability bounds.

To estimate the probability of the event (i), fix $t > t_0(\gamma, g)$ and set $(H_0, V_0) = (h, \nu) \in [0, tg/4\gamma(1+\gamma)] \times (a,b)$. System equations (2.1.1) show $H_u - V_u = h - \nu + (1+\gamma)S_u + ug - B_u$. For $u < \tau_a^V \wedge \tau_b^V \wedge t$,

$$\begin{aligned}
H_u - tg/2\gamma &= H_u - V_u + V_u - tg/2\gamma = h - \nu + (1+\gamma)S_u + ug - B_u + V_u - tg/2\gamma \\
&\leq tg/4\gamma(1+\gamma) + b - a + ug/2(1+\gamma) - tg/2\gamma + \sup_{u \leq t}(-B_u) \\
&\leq tg/4\gamma(1+\gamma) + b - a + t\left(g/2(1+\gamma) - g/2\gamma\right) + \sup_{u \leq t}(-B_u) = b - a - tg/4\gamma(1+\gamma) + \sup_{u \leq t}(-B_u),
\end{aligned} \quad (.0.12)$$

where the first inequality follows from $S_u \leq ub = -u\left(\frac{g - g/2(1+\gamma)}{1+\gamma}\right)$. If $\tau_a^V < \tau_b^V$ and $\tau_a^V \wedge \tau_b^V \leq t$, (.0.12) shows that if $H_{\tau_a^V} > tg/2\gamma + b - a$ then $\sup_{q \leq t}(-B_q) > tg/4\gamma(1+\gamma)$. Using (.0.12), a standard upper bound on the Gaussian distribution and the reflection principle, we obtain

$$\begin{aligned}
&\sup_{(h,\nu) \in [0, tg/4\gamma(1+\gamma)] \times (a,b)} \mathbb{P}_{(h,\nu)}\left(\tau_a^V < \tau_b^V,\, \tau_a^V \wedge \tau_b^V \leq t,\, H_{\tau_a^V} > tg/2\gamma + b - a\right) \\
&\qquad \leq \mathbb{P}\left(\sup_{u \leq t}(-B_u) > tg/4\gamma(1+\gamma)\right) \leq \frac{8(1+\gamma)\gamma}{g\sqrt{2\pi t}} e^{-tg^2/32\gamma^2(1+\gamma)^2}. \quad (.0.13)
\end{aligned}$$

To estimate the probability of event (ii), fix $t > t_0(\gamma, g)$ and set $(H_0, V_0) = (h, a)$ with any $h \in (0, tg/2\gamma + b - a]$. When $t < \sigma(0)$, (2.1.4) implies $0 < H_t \le tg/2\gamma + b - a + a/\gamma + g/\gamma^2 - B_t - tg/\gamma = c' - B_t - tg/2\gamma$ for a positive constant c'. We choose $t_0(\gamma, g)$ large enough that $t_0(\gamma, g)g/2\gamma - c' > 0$. Therefore,

$$\sup_{h\in(0,tg/2\gamma+b-a]} \mathbb{P}_{(h,a)}\left(t < \sigma(0)\right) \le \mathbb{P}\left(-B_t > tg/2\gamma - c'\right) \le \frac{\sqrt{t}}{\sqrt{2\pi}\,(tg/2\gamma - c')} e^{-(tg/2\gamma - c')^2/2t}. \tag{.0.14}$$

In (.0.11), (.0.13) and (.0.14) the choice of $t > t_0(\gamma, g)$ was arbitrary. As a result, (.0.11) and (.0.13) show that for a positive constant c,

$$\begin{aligned}
&\mathbb{P}_{(h,\nu)}\left(\sigma(\tau_a^V) - \tau_a^V > t,\ \tau_a^V < \tau_b^V\right) \\
&\qquad \le \mathbb{P}_{(h,\nu)}\left(t < \tau_a^V \wedge \tau_b^V\right) + \mathbb{P}_{(h,\nu)}\left(\tau_a^V < \tau_b^V,\ \tau_a^V \wedge \tau_b^V \le t,\ H_{\tau_a^V} > tg/2\gamma + b - a\right) \\
&\qquad\quad + \mathbb{P}_{(h,\nu)}\left(\sigma(\tau_a^V) - \tau_a^V > t,\ H_{\tau_a^V} \le tg/2\gamma + b - a,\ \tau_a^V < \tau_b^V,\ \tau_a^V \wedge \tau_b^V \le t\right) \\
&\qquad \le e^{-ct} + \mathbb{P}_{(h,\nu)}\left(\sigma(\tau_a^V) - \tau_a^V > t,\ H_{\tau_a^V} \le tg/2\gamma + b - a\right),
\end{aligned} \tag{.0.15}$$

holds for all $(h, \nu) \in [0, tg/4\gamma(1+\gamma)] \times (a, b)$ and t sufficiently large. The strong Markov property at τ_a^V and (.0.14) show

$$\sup_{(h,\nu)\in[0,tg/4\gamma(1+\gamma)]\times(a,b)} \mathbb{P}_{(h,\nu)}\left(\sigma(\tau_a^V) - \tau_a^V > t,\ H_{\tau_a^V} \le tg/2\gamma + b - a\right) \le e^{-ct}. \tag{.0.16}$$

(.0.15) and (.0.16) prove the lemma. □

Lemma .0.5. *For any $u \in \left(-\frac{g}{1+\gamma}, 0\right)$, there exists a constant $t_0(u, \gamma, g) > 0$ such that*

$$\sup_{\nu\in(u,0]} \mathbb{P}_{(0,\nu)}\left(\tau_u^V > t\right) \le \exp\left\{-\frac{\left(t\left((1+\gamma)u + g\right) + u - \nu\right)^2}{2t}\right\},$$

for $t > t_0(u, \gamma, g)$.

Proof. For $t < \tau_u^V$, $L_t \le \sup_{s\le t}(B_s - su) \le \sup_{s\le t} B_s - tu$ and

$$\sup_{s\le t} B_s - t(1+\gamma)u - tg + \nu \ge L_t - \int_0^t (\gamma V_s + g)\, ds + \nu = V_t > u.$$

The result follows by tail bounds on the supremum of Brownian motion, taking t large enough that $0 < \frac{\sqrt{t}}{t((1+\gamma)u+g)+u} < 1$. $\square$

The next lemma shows that the velocity process starting from a value $\nu \in (-\frac{g}{\gamma}, -\frac{g}{1+\gamma})$ cannot take too long to reach a larger value in the same interval because it is bounded below by a Brownian motion with positive drift.

Lemma .0.6. *Fix $\epsilon_0 \in (0, g/\gamma)$ and $u = -\frac{g+\epsilon_0}{1+\gamma}$. For each $h \geq 0$, there exists a constant $c > 0$ such that*

$$\sup_{\nu \in (-\frac{g}{\gamma}, u)} \mathbb{P}_{(h,\nu)}\left(\tau_u^V > t\right) \leq e^c \, e^{\epsilon_0 h} \, e^{-\frac{\epsilon_0^2}{2} t},$$

for all $t > t_0 := 2h/\epsilon_0 + 2g/\epsilon_0\gamma(1+\gamma) + 4/\epsilon_0^2$. In particular, there exists a constant $c' > 0$ depending on g, γ, ϵ_0 such that for any $\eta \in (0, \epsilon_0^2/4)$,

$$\sup_{\nu \in (-\frac{g}{\gamma}, u)} \mathbb{E}_{(h,\nu)} \, e^{\eta \tau_u^V} \leq e^{c'} e^{2\eta h/\epsilon_0}.$$

Proof. For $(H_0, V_0) = (h, \nu)$ with $h \geq 0, \nu \in \left(-\frac{g}{\gamma}, u\right)$, and any $t < \tau_u^V$,

$$V_t \geq -t(\gamma u+g)+\sup_{s\leq t}(-h+B_s-su)+\nu \geq B_t-t((1+\gamma)u+g)+\nu-h \geq B_t-t((1+\gamma)u+g)-\frac{g}{\gamma}-h.$$

This lower bound on V implies

$$\left\{\tau_u^V > t,\ (H_0, V_0) = (h,\nu)\right\} \subset \left\{\inf\left\{s \geq 0\ :\ B_s - s((1+\gamma)u+g) = u + \frac{g}{\gamma} + h\right\} > t\right\}.$$

From tail bounds on hitting times of Brownian motion with drift $-((1+\gamma)u+g) = \epsilon_0$, we conclude that

$$\begin{aligned}
\mathbb{P}_{(h,\nu)}\left(\tau_u^V > t\right) &\leq \int_t^\infty \frac{h+u+\frac{g}{\gamma}}{\sqrt{2\pi\, s^3}} \exp\left\{-\frac{\left(h+u+\frac{g}{\gamma}-\epsilon_0 s\right)^2}{2s}\right\} ds \\
&\leq \left(\frac{\epsilon_0}{2\theta}t + \frac{1}{2}\right)^{-1} \frac{1}{\sqrt{2\pi}} \int_{z(t)}^\infty \exp\left\{-\frac{z^2}{2}\right\} dz \\
&\leq \left(\frac{\epsilon_0}{\theta}t + 1\right)^{-1} \frac{1}{z(t)} \exp\left\{-\frac{z(t)^2}{2}\right\} \leq e^{\epsilon_0 \theta} \exp\left\{-\frac{\epsilon_0^2}{2}t\right\},
\end{aligned}$$

holds for all $\nu \in (-g/\gamma, u)$ and for $t \geq 2\left(\frac{\theta}{\epsilon_0} \vee \frac{2}{\epsilon_0^2}\right)$, where $\theta = h + g/\gamma + u > 0$. The second inequality follows after a change of variables to $z(s) = \epsilon_0\sqrt{s} - \frac{\theta}{\sqrt{s}}$. t_0 is chosen to satisfy $t_0 \geq 2\left(\frac{\theta}{\epsilon_0} \vee \frac{2}{\epsilon_0^2}\right)$ and $z(t) \geq \epsilon_0\sqrt{t} - \frac{h+g/\gamma - g/(1+\gamma)}{\sqrt{t}} = \epsilon_0\sqrt{t} - \frac{h+g/\gamma(1+\gamma)}{\sqrt{t}} \geq 1$ for $t \geq t_0$.

The second statement follows from $\sup_{\nu\in(-\frac{g}{\gamma},u)} \mathbb{E}_{(h,\nu)}\, e^{\eta\,\tau_u^V} \leq \int_0^\infty \sup_{\nu\in(-\frac{g}{\gamma},u)} \mathbb{P}_{(h,\nu)}\left(e^{\eta\,\tau_u^V} > t\right) dt$. $\square$

Remark .0.1. *Writing $\tau_u^{V+} = \inf\{t > 0 : V_t = u\}$, the proof of Lemma .0.6 shows that the same bounds obtained above hold for $\mathbb{P}_{(h,u)}\left(\tau_u^{V+} > t\right)$ and $\mathbb{E}_{(h,u)}\, e^{\eta\tau_u^{V+}}$ when $h > 0$.*

Lemma .0.7. *Fix $u > 0$. For each $\nu > u$,*

$$\sup_{h\geq 0} \mathbb{P}_{(h,\nu)}\left(\tau_u^V > t\right) \leq 1 \wedge \exp\left\{-2u\left(u - \nu + t(\gamma u + g)\right)\right\},$$

for all $t \geq 0$. Therefore, if $\eta \in (0, u(\gamma u + g))$

$$\sup_{h\geq 0} \mathbb{E}_{(h,\nu)}\, e^{\eta\,\tau_u^V} \leq e^{\frac{\eta}{\gamma u+g}(\nu-u)} + e^{2u(\nu-u)-\frac{\eta}{\gamma u+g}(\nu-u)} \leq 2e^{2u(\nu-u)}.$$

Proof. Fix $(H_0, V_0) = (h, \nu) \in \mathbb{R}_+ \times (u, \infty)$. The definition of V gives for $t < \tau_u^V$,

$$u < V_t \leq \nu - t(\gamma u + g) + 0 \vee \sup_{s\leq t}\left(-h + B_s - su\right) \leq \nu - t(\gamma u + g) + \sup_{s<\infty}\left(B_s - su\right).$$

This bound implies the first claim of the lemma upon using the fact that for $u > 0$, $\sup_{s<\infty}(B_s - su) \stackrel{d}{=} \text{Exponential}(2u)$ (see Chapter 3.5 of Karatzas and Shreve (1991)). The second claim is a consequence of the first. $\square$

BIBLIOGRAPHY

Aldous, D. and Diaconis, P. (1986). Shuffling cards and stopping times. *American Mathematical Monthly*, 93(5):333–348.

Alfonsi, A., Jourdain, B., and Kohatsu-Higa, A. (2014). Pathwise optimal transport bounds between a one-dimensional diffusion and its Euler scheme. *The Annals of Applied Probability*, 24(3):1049–1080.

Andres, S. (2009). Diffusion processes with reflection. *PhD thesis*, TU Berlin.

Arbabi, H. and Sapsis, T. (2020). Generative stochastic modeling of strongly nonlinear flows with non-Gaussian statistics. *arXiv preprint, https://arxiv.org/pdf/1908.08941.pdf.*

Atar, R., Budhiraja, A., and Dupuis, P. (2001). On positive recurrence of constrained diffusion processes. *Ann. Probab.*, 29(2):979–1000.

Athreya, A. et al. (2012). Propagating Lyapunov functions to prove noise-induced stabilization. *Electronic Journal of Probability*, 17(96):1–38.

Bakry, D., Gentil, I., and Ledoux, M. (2014). *Analysis and geometry of Markov diffusion operators*. Springer International Publishing Switzerland.

Banerjee, S. and Budhiraja, A. (2020). Parameter and dimension dependence of convergence rates to stationarity for reflecting Brownian motions. *The Annals of Applied Probability*, 30(5):2005–2029.

Banerjee, S. and Burdzy, K. (2020). Rates of convergence to equilibrium for Potlatch and Smoothing processes. *The Annals of Probability,* to appear.

Banerjee, S., Burdzy, K., and Duarte, M. (2019). Gravitation versus Brownian motion. *Annales de L'Institut Henri Poincaré, Probabilités et Statistiques*, 55(3):1531–1565.

Banerjee, S. and Kendall, W. (2016). Coupling the Kolmogorov diffusion: Maximality and efficiency conditions. *Advances in Applied Probability*, 48(A):15–35.

Banerjee, S. and Mukherjee, D. (2019a). Join-the-shortest queue diffusion limit in halfin–whitt regime: Tail asymptotics and scaling of extrema. *The Annals of Applied Probability*, 29(2):1262–1309.

Banerjee, S. and Mukherjee, D. (2019b). Join-the-shortest queue diffusion limit in halfin-whitt regime: Sensitivity on the heavy-traffic parameter. *To appear in The Annals of Applied Probability.*

Barnes, C. (2018). Convergence of jump processes with stochastic intensity to Brownian motion with inert drift. *arXiv: 1109.04428.*

Bass, R., Burdzy, K., Chen, Z.-Q., and Hairer, M. (2010). Stationary distributions for diffusions with inert drift. *Probability Theory and Related Fields*, (146):1–47.

Bertoin, J. (2019). On a Feynman-Kac approach to growth-fragmentation semigroups and their asymptotic behaviors. *Journal of Functional Analysis*, 277(11).

Bion-Nadal, J. and Talay, D. (2019). On a Wasserstein-type distance between solutions to stochastic differential equations. *The Ann. of Appl. Probab.*, 29(3):1609–1639.

Biswas, N., Jacob, P., and Vanetti, P. (2019). Estimating convergence of markov chains with l-lag couplings. *33rd Conference on Neural Information Processing Systems.*

Blanchet, J. and Chen, X. (2020). Rates of convergence to stationarity for reflected Brownian motion. *Mathematics of Operations Research*, 45(2):660–681.

Blanchet, J., Chen, X., Glynn, P., and Si, N. (2020). Efficient steady-state simulation of high-dimensional stochastic networks. *arXiv preprint arXiv:2001.08384*.

Budhiraja, A. and Lee, C. (2007). Long time asymptotics for constrained diffusions in polyhedral domains. *Stochastic Processes and their Applications*, 117(2):1014–1036.

Burdzy, K. and White, D. (2008). Markov processes with product-form stationary distribution. *Electron. Commun. Probab.*, 13:614–627.

Butkovsky, O. et al. (2020). Generalized couplings and ergodic rates for SPDEs and other Markov models. *The Ann. of Appl. Probab.*, 30(1):1–39.

Cattiaux, P. and Guillin, A. (2017). Hitting times, functional inequalities, Lyapunov conditions and uniform ergodicity. *Journal of Functional Analysis*, 272(6):2361–2391.

Cloez, B. and Hairer, M. (2015). Exponential ergodicity for Markov processes with random switching. *Bernoulli*, 21(1):505–536.

Cooke, B., Herzog, D., Mattingly, J., McKinley, S., and Schmidler, S. (2017). Geometric ergodicity of two-dimensional Hamiltonian systems with a Lennard-Jones-like repulsive potential. *Communications in Mathematical Sciences*, 15(7):1987–2025.

Dembo, A., Gantert, N., and Zeitouni, O. (2004). Large deviations for random walk in random environment with holding times. *The Annals of Probability*, 32(1B):996–1029.

Down, D., Meyn, S., and Tweedie, R. (1995). Exponential and uniform ergodicity of Markov processes. *The Annals of Probability*, 23(4):1671–1691.

Durrett, R. (2010). *Probability: Theory and Examples.* Cambridge University Press, 4th edition.

Eberle, A. (2016). Reflection couplings and contraction rates for diffusions. *Probab. Theory Relat. Fields*, (166):851–886.

Einstein, A. (1956). *Investigations on the theory of the Brownian movement.* Courier Corporation.

Émery, M. (2005). On certain almost Brownian filtrations. *Ann. Inst. H. Poincaré Probab. Statist.*, 43(3):285–305.

Eschenfeldt, P. and Gamarnik, D. (2018). Join the shortest queue with many servers: The heavy-traffic asymptotics. *Mathematics of Operations Research*, 43(3):867–886.

Evans, L. (2010). *Partial Differential Equations.* American Mathematical Society, 2nd edition.

Fang, S., Shao, J., and Sturm, K.-T. (2010). Wasserstein space over the Wiener space. *Probab. Theory Relat. Fields*, (146):535–565.

Fernholz, E. R. (2002). Stochastic portfolio theory. In *Stochastic Portfolio Theory*, pages 1–24. Springer.

Feyel, D. and Üstünel, A. (2004). Monge-Kantorovitch measure transportation and Monge-Ampère equation on Wiener space. *Probab. Theory Relat. Fields*, (128):347–385.

Flamary, R. and Courty, N. (2017). POT Python optimal transport library.

Friedman, A. (1975). *Stochastic Differential Equations and Applications, volume I.* Academic Press.

Hairer, M. and Mattingly, J. (2011). Yet another look at Harris' ergodic theorem for Markov chains. *Seminar of Stochastic Analysis, Random Fields and Applications VI*, pages 109–117.

Harrison, J. and Reiman, M. (1981). Reflected Brownian motion on an orthant. *Ann. Probab.*, (9):302–308.

Harrison, J. and Williams, R. (1987a). Brownian models of open queueing networks with homogeneous customer populations. *Stochastics*, 22(2):77–115.

Harrison, J. and Williams, R. (1987b). Multidimensional reflected Brownian motions having exponential stationary distributions. *The Annals of Probability*, 15(1):115–137.

Herzog, D. and Mattingly, J. (2015). A practical criterion for positivity of transition densities. *Nonlinearity*, 28(8):2823–2825.

Hoff, P. (2009). *A first couple in Bayesian statistical methods.* Springer.

Ichiba, T., Karatzas, I., and Shkolnikov, M. (2013a). Strong solutions of stochastic equations with rank-based coefficients. *Probab. Theory Relat. Fields*, 156:229–248.

Ichiba, T., Pal, S., and Shkolnikov, M. (2013b). Convergence rates for rank-based models with applications to portfolio theory. *Probab. Theory Relat. Fields*, (156):415–448.

Ichiba, T., Papathanakos, V., Banner, A., Karatzas, I., and Fernholz, R. (2011). Hybrid Atlas models. *The Annals of Applied Probability*, 21(2):609–644.

Ikeda, N. and Watanabe, S. (1989). *Stochastic Differential Equations and Diffusion Processes.* North-Holland Publishing.

Jourdain, B. and Malrieu, F. (2008). Propagation of chaos and Poincaré inequalities for a system of particles interacting through their cdf. *The Annals of Applied Probability*, 18(5):1706–1736.

Jourdain, B. and Reygner, J. (2013). Propagation of chaos for rank-based interacting diffusions and long time behaviour of a scalar quasilinear parabolic equation. *Stochastic partial differential equations: analysis and computations*, 1(3):455–506.

Karatzas, I. and Shreve, S. (1991). *Brownian Motion and Stochastic Calculus.* Springer-Verlag, 2nd edition.

Karatzas, I. and Pal, S. and Shkolnikov, M. (2016). Systems of Brownian particles with asymmetric collisions. *Annales de l'I.H.P. Probabilités et statistiques*, 52(1):323–354.

Kella, O. and Ramasubramanian, S. (2012). Asymptotic irrelevance of initial conditions for Skorohod reflection mapping on the nonnegative orthant. *Mathematics of Operations Research.*

Kendall, W. (2009). Brownian couplings, convexity, and shy-ness. *Electron.Commun.Probab.*, 14:66–80.

Kersaudy, P. et al. (2015). A new surrogate modelingtechnique combining Kriging and polynomial chaos expansion. *Journal of Computational Physics*, (286):130–117.

Knight, F. B. (2001). On the path of an inert particle impinged on one side by a Brownian particle. *Probability Theory and Related Fields*, 121:577–598.

Lax, P. (2002). *Functional Analysis.* Wiley-Interscience.

Le Gall, J.-F. (2013). *Brownian Motion, Martingales and Stochastic Calculus (English).* Springer-Verlag Berlin.

Levin, D. and Peres, Y. and Wilmer, E. (2017). *Markov Chains and Mixing Times.* American Mathematical Society, 2nd edition.

Lions, P.-L. and Sznitman, A.-S. (1984). Stochastic differential equations with reflecting boundary conditions. *Comm. Pure Appl. Math.*, 37:511–537.

Mandelbaum, A. and Ramanan, K. (2010). Directional Derivatives of Oblique Reflection Maps. *Mathematics of Operations Research*, 35(3):527–558.

Mattingly, J., McKinley, S., and Pillai, N. (2012). Geometric ergodicity of bead-spring pair with stochastic stokes forcing. *Stochastic Processes and their Applications*, (122):3953–3979.

Mattingly, J. and Stuart, A. (2002). Geometric ergodicity of some hypo-elliptic diffusions for particle motions. *Markov Processes and Related Fields*, 8(2):199–214.

Meyn, S. and Tweedie, R. (1993a). Stability of markovian processes II: Continuous-time processes and sampled chains. *Advances in Applied Probability*, 25(3):487–517.

Meyn, S. and Tweedie, R. (1993b). Stability of markovian processes III: Foster-lyapunov criteria for continuous-time processes. *Advances in Applied Probability*, 25(3):518–548.

Meyn, S. and Tweedie, R. (2009). *Markov Chains and Stochastic Stability.* Cambridge University Press, 2nd edition.

Meyn, S. and Tweedie, R. (2012). *Markov chains and stochastic stability.* Springer Science & Business Media.

Nualart, D. (2006). *The Malliavin calculus and related topics.* Springer-Verlag.

O'Connor, K., McGoff, K., and Nobel, A. (2020). Optimal transport for stationary Markov chains via policy iteration. *arXiv preprint, https://arxiv.org/pdf/2006.07998.pdf.*

Pal, S. and Pitman, J. (2008). One-dimensional Brownian particle systems with rank-dependent drifts. *The Annals of Applied Probability*, 18(6):2179–2207.

Pal, S. and Sarantsev, A. (2019). A note on transportation cost inequalities for diffusions with reflections. *Electronic Communications in Probability*, 24.

Peyré, G. and Cuturi, M. (2018). *Computational Optimal Transport.* ArXiv:1803.00567.

Pipiras, V. and Brown, B. (2019). On extending multifidelity uncertainty quantification methods from non-rare to rare problems. *Proceedings of the 17th International Ship Stability Workshop.*

Pitman, J. (1974). Uniform Rates of Convergence for Markov Chain Transition Probabilities. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, (29):193–227.

Ramasubramanian, S. (1996). Transition densities of reflecting diffusions. *Sankhya*, 58(3):347–381.

Ramsay, J. and Silverman, B. (2005). *Functional Data Analysis.* Springer, 2nd edition.

Reiman, O. (1984). Open Queueing Networks in Heavy Traffic. *Mathematics of Operations Research*, 9(3):441–458.

Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion.* Springer-Verlag, 3rd edition.

Santambrogio, F. (2015). *Optimal Transport for Applied Mathematicians.* Birkhäuser.

Sarantsev, A. (2015). Triple and simultaneous collisions of competing Brownian particles. *Electronic Journal of Probability*, 20.

Sarantsev, A. and Tsai, L.-C. (2017). Stationary gap distributions for infinite systems of competing Brownian particles. *Electronic Journal of Probability*, 22.

Sarantsev, A. (2017a). Infinite Systems of Competing Brownian Particles. *Annales de l'I.H.P. Probabilités et statistiques*, 53(4):2279–2315.

Sarantsev, A. (2017b). Reflected Brownian Motion in a Convex Polyhedral Cone: Tail Estimates for the Stationary Distribution. *Journal of Theoretical Probability*, 30:1200–1223.

Sarantsev, A. (2019). Comparison Techniques for Competing Brownian Particles. *Journal of Theoretical Probability*, 32:545–585.

Simon, B. (2015). *Real Analysis: A comprehensive course in analysis, part 1.* American Mathematical Society.

Soize, C. and Ghanem, R. (2004). Physical systems with random uncertainties: Chaos representations with arbitrary probability measure. *SIAM J. Sci. Comput.*, 26(2):395–410.

Stroock, D. (2008). *Partial differential equations for for probabilists.* Cambridge University Press.

Sznitman, A.-S. (2004). *Topics in Random Walks in Random Environment.* School and Conference on Probability Theory, ICTP Lecture Notes Series.

Tang, W. (2019). Exponential ergodicity and convergence for generalized reflected Brownian motion. *Queueing Systems*, 92:83–101.

Thorisson, H. (2000). *Coupling, Stationarity and Regeneration.* Springer-Verlag New York.

Villani, C. (2009). Hypocoercivity. *Memoirs of the American Mathematical Society*, (950).

Villani, C. (2009). *Optimal Transport: Old and New*. Springer-Verlag.

White, D. (2007). Processes with inert drift. *Electronic Journal of Probability*, 12(55):1509.

Willcox, K. (2020). Scientific machine learning: Where physics-based modeling meets data-driven learning. *Talk at JuliaCon, https://www.youtube.com/watch?v=Bk4PJnjuPps*.

Xiu, D. and Karniadakis, G. (2002). The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24(2):619–644.